International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 06 | June -2018 www.irjet.net p-ISSN: 2395-0072

Real Time Hand Gesture Recognition Using Finger Tips

Shivendra Singh

Student, Dept. of Electronic and Communication Engineering, Vellore Institute of Technology, Chennai Campus, India

---

Abstract – *Over the years human-computer interaction has evolved in many ways. Being an important part of non-verbal communication which plays a vital role in our daily life, hand gesture controlled man-machine interaction is one of the most significant one. This paper presents a novel real-time method for hand gesture recognition. In this framework, the hand region is extracted from the background by skin color threshold. The palm center and finger tips are located after segmentation of finger-palm region. The segmented region is centered and the angle between finger tips and palm center is calculated. A rule classifier makes use of the number of finger tips detected and the calculated angle to predict the label of hand gesture*

Key Words: Gesture Recognition, HCI, MMI, Classifier, Segmentation, Threshold, HSI, Distance transform

1. INTRODUCTION

The essential aim of building a hand gesture recognition system is to eliminate the need of any physical contact between the controller and the controlled object in any way. Hand gesture has a natural ability to represent ideas and actions which humans have been using for thousands of years. Being able to classify and interpret these gestures can provide a more natural interaction to the computer system. This type of interaction will be the core of the augmented reality. Hand gesture also has a great value in many applications such as sign language recognition to communicate with deaf and mute person, Wireless robotic control [1], Virtual reality and augmented reality. The first way of interaction with computer using hand gesture was proposed by Myron W.Krueger in 1970 [2]. Recent advancement has forced developers to explore new method of HCI which are faster and are have robust recognition than ever.

Gesture is a physical indication of the body to convey information. Though any bodily movement can be considered a gesture, generally it originates from the movement of hand or face or combination of both. Combined gesture are quiet complex and difficult for a machine to classify. For example, waving the hand is a gesture that can suggest “hi” or “bye” depending upon the facial expression, differentiating this gesture requires an emotional understanding (which is still under investigation). Simple gestures can be classified as Static and Dynamic also termed as micro gestures and macro gestures respectively. For the former, the posture defines the expression. Movements are the defining entity for the latter one.

The key problem in gesture interaction is how to make a computer understood these hand gestures. Over the years researchers have proposed different techniques to train the system to identify these gestures with more accuracy.

This paper discusses static gesture recognition. The overall aim of this paper is to make the computer understand human gesture and differentiate them. The algorithm is trained to classify five different hand gestures.

The paper is organized as follow. Section 2 deals with related work. Section 3 is the proposed work that include 5 stages namely hand detection, pre-processing, palm point detection, finger tips detection and hand gesture recognition. Section 4 deals with results and discussion followed by conclusion in Section 5

2. RELATED WORK

Zhi-hua Chen et al [3] in their paper presented a real time method for hand gesture recognition. The framework first extracts the hand region from the background using background subtraction method. The output of the hand detection is a binary image in which the white pixels are members of the hand region, while the black pixels belong to background region. Palm point, center of the palm is used to create a palm mask. This palm mask is used to segment fingers and palm. The method is rotation invariant as it uses palm point and wrist point to align the segmented image. Thumb and fingers are detected and a rule classifier is applied to predict the label of hand gesture.

Hsiang-Yueh Lai recognized eleven kinds of hand gesture using convex defect character points [4]. The hand is extracted form the background using YCbCr color transformation and hand contours are defined. These hand contours are used to get conves defect character points, and finger angle and fingertip positions are calculated to recognize the hand gesture. Thang B. Dinh used boosted cascade of classifiers to recognize hand gestures [5]. Their system is able to efficiently recognize 24 basic signs of American Sign Language (ASL). Use of AdaBoost and informative Haar wavelet features helped in achieving high computational performance. 24 detectors are trained for 24 gestures with an average number of stages equal to 13. The gesture AEST got a big false alarm because of similarity between them. The system has a correct rate of 83%. Freeman used Orientation Histogram to represent hand posture [6] which has low computational time, but requires the input to be close-up image of the pattern.

Moreover, Orientation Histogram is similar for signs that are very similar to one another, which makes the classifier inaccurate. Ahmed et al. [7] introduced the use of a Bayesian

technique in order to recognize 3D hand gestures even with large amounts of noise and uncertain or missing input data.

3. PROPOSED WORK

Figure-1 gives a workflow for the proposed hand gesture recognition system. First the hand is detected using HSI (Hue, Saturation, and Intensity) values of the skin. The result is pre-processed to remove noise and non-hand region. The output of the pre-processing is converted into a binary image. The palm point (center of palm) and finger tips are detected using binary image. From the palm point and finger tips the angle between fingers is calculated and a simple classifier rule is used to recognize the hand gesture.

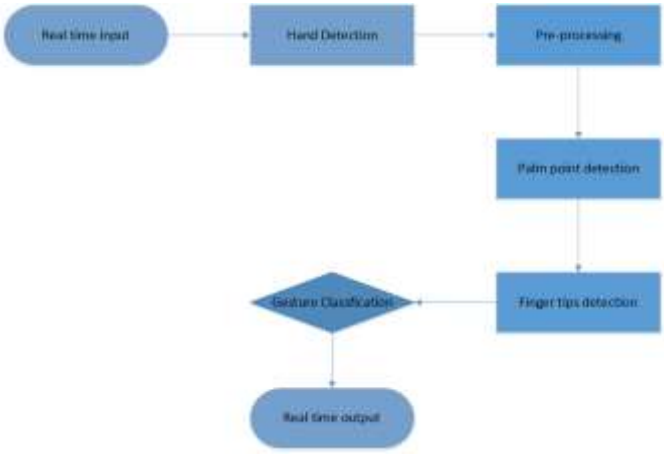


Figure-1: The workflow of the proposed method

3.1 Hand Detection

Figure-2 shows an original image used for hand gesture recognition in this work. These images are captured using a normal camera and all the images are taken under identical condition i.e. controlled illumination. So, it is easy and effective to detect the hand region. The color of the skin is measure with the HSI model. The upper HSI values of skin color are taken as 20, 255 and 255 respectively. The lower HSI values of the skin color are taken as 0, 50, and 90 respectively. A filter is used to filter out other values that do not fall in this range.

3.2 Pre-processing

The detected hand region contains arm region with other noises due to color similarity in the background. The detected region is converted into binary image. The region which represents hand has all the pixel values set as 0 and everything else is considered background with pixel intensity set as 1. To reduce the background noise we remove all small insignificant smudges or connected components from the images those have area fewer than P pixels. To segment the hand, boundary contour is calculated in the image.

It is performed by scanning the whole image from left to right and then in that scanned boundary from top to bottom. Area and length of the contour are used to segment palm-finger region from arm region. The output after hand segmentation and noise removal is shown in figure 3.

Figure -2: Original input image

Figure-3: Output after hand segmentation and noise removal

3.3 Palm point detection

The palm point is defined as the center of the palm. Distance transform is used to find the center of the palm. Also called as distance map, distance transform is the representation of image according to distance of pixels. In distance transform image, each pixel records the distance of it and the nearest boundary pixel. An example of distance transform is shown in figure 4. In figure 4(a) a binary image is shown. Figure 4(b) shows the distance transform of the figure 4(a).

The method used to measure the distance between the pixels and the nearest boundary pixels is City Block distance. As shown in figure 4(b), the center point of the binary image is with the largest distance four. Thus, in the distance transform image of the binary hand image, the pixel with the largest distance is chosen as the palm point. The found palm point is marked in the distance transformed image shown in figure 5. The image is resized and centered to make the process scale invariant.

0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0

(a)

0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0
0	1	2	2	2	2	2	1	0
0	1	2	3	3	3	2	1	0
0	1	2	3	4	3	2	1	0
0	1	2	3	3	3	2	1	0
0	1	2	2	2	2	2	1	0
0	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0

(b)

Figure-4: An example of distance transform; **(a)** binary image; **(b)** distance transform of (a).

3.3.1 City block distance

Also known as Manhattan distance, the city block distance is calculated as the distance in x plus the distance in y. It is the absolute difference between coordinates of pair of objects. This distance is always greater than or equal to zero.

$$CityBlock_{Distance} = \sum_{j=1}^{k} |a_j - b_j|$$

Where k is the dimension and a, b are two points, $CityBlock_{Distance}$ is the distance between point 'a' and 'b'.

3.4 Finger-tip detection

In this step, tip of the finger is denoted as peak. Total numbers of fingers raised are found by number of peaks detected. The fact that fingers are of different height is applied. To find the finger tips, the hand region is divided into two vertical regions. The image is scanned for each halves. First half is scanned from left to right, starting from top and second half from right to left starting from top. Whenever a pixel of '0' intensity is detected (hand region pixel value set to 0 in pre-processing), it is marked as tip. The vertical half in which the tip is detected is reduced to new size (where the tip was detected); similarly the top-most point is also reset. Again the image is scanned for new halves and this process continues.

Some false peak may be detected during fingertip detection. These peaks can result in false classification. To eliminate these unwanted peaks a threshold is computed. Here 0.07 times the maximum peak value is used as threshold. Values greater than this threshold are considered as peak.

3.4.1 Euclidean distance

The Euclidean distance between the points P and Q is the length between the line segments joining them.

$$Euclidean_{distance} = \sqrt{(x_1 - x_2)^2 - (y_1 - y_2)^2}$$

Where (x_1, y_1) are coordinate of point P and (x_2, y_2) are coordinate of point Q. $Euclidean_{distance}$ is the distance between point P and Q.

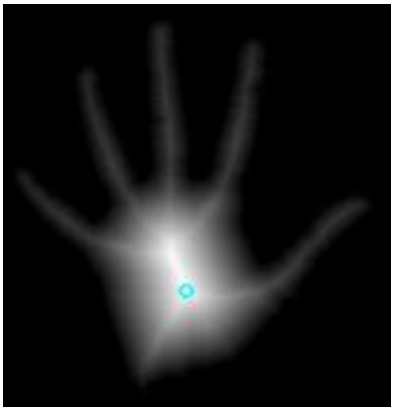


Figure-5: Distance transform of hand image with palm point marked on it.

3.5 Gesture classification

Classification of various hand gestures is based on numbers of finger tips detected and the angle between each fingertip and the palm point. For example, if three fingers are detected, and the angle formed by each finger with the palm point is a1, a2, a3 (matching the training database) only then it will be labeled as three.

3.5.1 Angle between two points

The angle between two points P and Q is calculated along the horizontal axis.

$$Angle = tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right)\left(\frac{180}{\pi}\right)$$

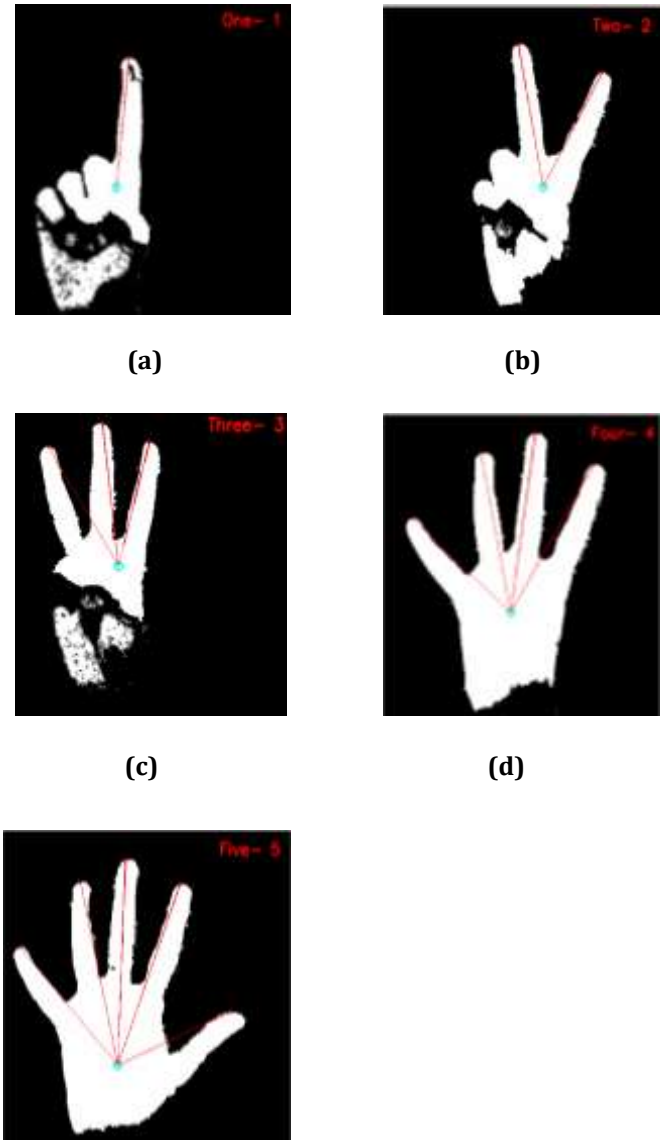
Where (x_1, y_1) are coordinate of point P and (x_2, y_2) are coordinate of point Q.

4. RESULTS AND DISCUSSION

There are some constrains and conditions assumed in the proposed work for better results. First is the generic environment- The proposed method works well; without static or homogeneous background. The system relies on hue, saturation, and intensity values to detect the hand region, a variation in light intensity can affect the hand detection rate. Hence, all training and testing phase was

performed in controlled environment; fixed light intensity was used.

The proposed algorithm can work with regular webcam camera, which is used to capture live video of the users hand gesture. The proposed work is able to recognize five hand gestures shown in figure 6. In the training phase the angle between different fingers with the palm point is calculated and it is associated with the number of finger tips found. The method was tested using hand gestures at different scale i.e. the input was given at different distance from the camera. This ensured as if the algorithm was tested on different person with varying hand size. Figure 7 shows a gesture with two finger raised but the angle formed by the tip of fingers with palm point is different from the trained data. Hence it is marked as not defined. Table 1 shows hit rate of test cases and figure 8 is the diagram based on table 1



(a) **(b)** **(c)** **(d)** **(e)**

Figure-6: Five classified hand gestures **(a)** One **(b)** Two **(c)** Three **(d)** Four **(e)** Five

Figure-7: Hand gesture not found in training dataset.

Gesture	Test Case	Match	Mismatch	Hit rate
One	120	110	10	91.6 %
Two	120	112	8	93.3 %
Three	120	103	17	85.8 %
Four	120	117	3	97.5 %
Five	120	120	0	100 %

Table-1: Recognition rate of each gesture.

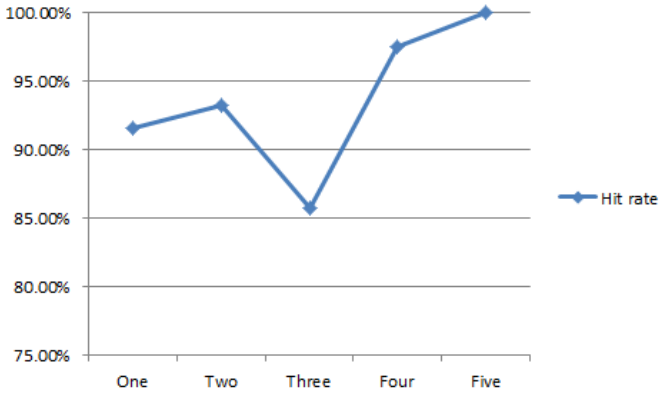


Figure-8: plot based on table 1.

5. CONCLUSION AND FUTURE WORK

The proposed work introduces a hand gesture classification system that is able to classify five different gestures. The hand region is detected form the background using HSI skin color model. The noise is removed from the detected hand and palm point and finger tips are detected. The recognition of hand gesture is accomplished by a rule classifier based on angles between palm point and finger tips and number of finger tips. The experimental result shows that the approach performs well and is fit for real time application.

The performance of the proposed method highly depends on the result of hand detection. As HSI skin color model is used to detect hand, the method is susceptible to light intensity. Also if objects of considerable size with color similar to skin are present in background, the performance of the system degrades.

However machine learning algorithms can differentiate the hand from the background. In future works, machine learning method may be used to address the complex background problem.

REFERENCES

[1] G. Chanhan and P. Chandhari, "Gestures based wireless robotic control using image processing," *2015 5th* Nirma University International Conference on Engineering (NUiCONE), Ahmedabad, 2015, pp. 1-7.

[2] M. KRUEGER. 1991. Artificial reality II, Addison-Wesley Reading.

[3] Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang, and Yu-Bo Yuan, "Real-Time Hand Gesture Recognition Using Finger Segmentation," The Scientific World Journal, vol. 2014, Article ID 267872, 9 pages, 2014.

[4] Hsiang Yueh Lai, Han Jheng Lai 2014. "Real-Time Dynamic Hand Gesture Recognition" in Computer, Consumer and Control (IS3C), 2014 International Symposium, IEEE

[5] Dinh, Thang & B. Dang, Van & Duc, Duong & Nguyen, Tuan & Le, Duy-Dinh. (2006). Hand gesture classification using boosted cascade of classifiers. Proceedings of the 4th IEEE International Conference on Research, Innovation and Vision for the Future, RIVF'06. 139 - 144. 10.1109/RIVF.2006.1696430.

[6] W, FREEMAN. C, WEISSMAN. 1995. Television control by hand gestures. Proc. of Intl. Workshop on Automatic Face and Gesture Recognition, 1995. 179-183.

[7] S. Ahmad, V. Tresp, "Classification With Missing And Uncertain Inputs", IEEE International Conference On Neural Network", 28 Mar. 1 Apr. 1993, Vol. 3, Pp. 1949 - 1954