Real Time Static and Dynamic Hand Gestures Cognizance for Human Computer Interaction

2021 10th International Conference on Advances in Computing and Communications (ICACC) | 978-1-6654-3919-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICACC-202152719.2021.9708249

Ajith Jacob
Department of Computer Science and Engineering,
Kottayam Institute of Technology and Science
Kottayam, India
ajithjacob23@gmail.com

Monish Koshy
Department of Mechanical Engineering
Kottayam Institute of Technology and Science
Kottayam, India
monishkoshy@tutanota.com

Nisha K. K
Department of Computer Science and Engineering
Rajiv Gandhi Institute of Technology
Kottayam, India
knisha81@gmail.com

***Abstract*—Human-Computer Interaction (HCI) has become ubiquitous. Hand gestures are integral part of HCI. Static and dynamic hand gestures used for HCI in real time systems are area of vital analysis, with manifold feasible applications. Vision based hand gesture recognition techniques had proven many advantages compared with hardware devices, as it provides end user an elementary and more perceptive method of communication between user and system. Although a wide range of recognition systems based on gesture is available, they are expensive and complex. Some existing gesture recognition methods are intended for usage with either statics or dynamic gestures. In this paper, a low-cost system for recognition of both static and dynamic hand gestures, based on computer vision is proposed. The application of vision based technology for managing the devices by gestures, dwindle the required work space.**

Keywords—Human Computer Interaction Applications, Hand Gestures Identification, Static and Dynamic Hand Recognition.

I. INTRODUCTION

HCI has become ubiquitous, it is typically achieved with the help of a physical controller like a joystick, touch screen or mouse. But these hardware devices hinder Natural User Interface (NUl) as there exits a strong boundary between the end user and system. Whereas hand gesture recognition methods based on vison have unparalleled advantages over conventional devices, as it provides the end user an elementary and more perceptive method of communication between the computer. Moreover, the visual data input makes it possible to connect remotely with computerized equipment, without any direct contact.

Hand gestures is also a part of HCI. The two classes of hand gestures used are static hand gestures and dynamic hand gestures. Some existing gesture recognition methods are intended for usage with either static or dynamic gestures. Overall, there are hardly any solutions that work synchronously on both static and dynamic gestures.

The aim of this research is to design an advanced, simple and real-time solution based on computer vision; for the identification of static and dynamic gestures. An elegant user interface is also implemented in this work. By applying vision based technology for managing the devices by gestures, the work space required can be dwindled. Gesture recognition has multiple applications in the context of sign language interpretation, smart home interactive control, control of a system interface, control of virtual environments, human robot interaction and robot hand control.

This paper is organized as follows. Following the introduction, Section II delve into the related works done in this field. Section III depicts the proposed system architecture. The experimentation results are briefed in Section IV. Finally, the conclusion of the work is presented in Section V.

II. RELATED WORKS

Implementation of static and dynamic gesture for HCI in real time system remains an active field of study, it has multiple application in the area of control of virtual environments, human robot interaction, robot hand control etc. Some of the pioneer research works in this area are explained below.

A. Static Hand Gestures

In 2011, Dipak Kumar Ghosh et al. [1] propounded a sagacious program algorithm for static hand movement identification which solves the challenges such as size, orientation and coordinate position variation of the images in hand gesture image recognition. RBF neural network which is primarily based on K-mean was used as classifier for identification of alphabets in sign language from static gesture image.

Trong-Nguyen Nguyen et al. [2] presented a low computational system to identify static gestures. In this system segmentation based on colour information and geometrical properties are used to analyze the detailed information of hand. Then for the identification a max-win voting strategy was implemented using SVM model.

B. Dynamic Hand Gestures

In 2011, Hong-xiang et al. [3] presented a sagacious model to identify motion trajectory of hand using initial posture and final posture of hand in space. Mean shift algorithm with is based on motion and colour details of hand is used to detect the trajectory of hand gesture.

Somayeh Shiravandi et al. [4] designed a hand gesture recognition using Bayesian dynamic network. This model operates on two equivalent networks and it showed an accuracy about 90%.

C. Both Static and Dynamic Hand Gestures

Guillaume Plouffe et al. [5] propounded a method for gesture identification in which Kinect sensor is used for data acquisition. Block search, directional search and *k*-curvature algorithm are used to recognize hand profile in space. Besides, dynamic time warping algorithm is used for identification of gestures.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed system solves the problem of static and dynamic hand recognition with a low-cost camera. The end user's hand movements captured using camera in live streaming is the data input for the system. The algorithm developed for identification follows three steps; i.e. hand area detection in space, features extraction of palm and gesture recognition.

A. *Hand Area Detection*

An algorithm based on skin color is applied to extracts the tract of hand. The identification of hand tract is mainly used for subtracting the background with respect to the lighting of the image. Different steps used in identification hand tract is shown in Fig. 1.

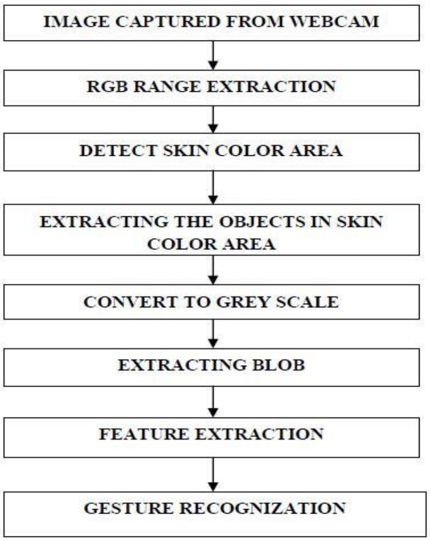

Fig. 1. Steps in hand area detection.

The image is captured from the webcam. After that the data acquisition is set to the RGB range, within a range of minimum and maximum range of RGB. From the minimum and maximum range of RGB, skin colour region is detected using Euclidean filter. The Euclidean filter filters pixels with respect to colors inside/outside of RGB sphere with specified centre and radius.

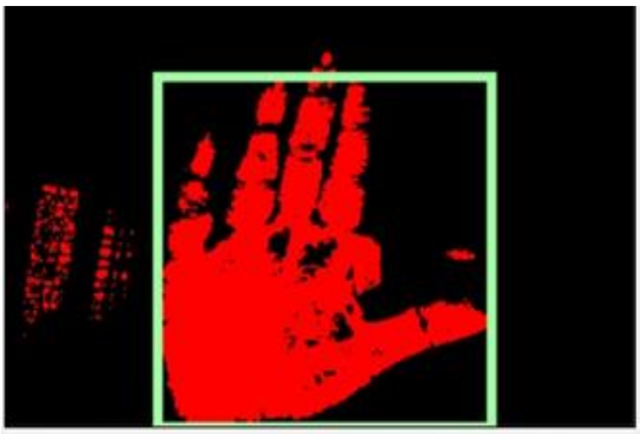

Fig. 2. Blob detection.

In Euclidean filtering, image is converted to grey scale image; it is part of an internal process. This image is converted to red- black image where skin part become red. From this red and black image, blob is extracted shown in Fig. 2. Blob detection methods are aimed at detecting areas in a digital image that differ in properties such as brightness or color, compared to neighboring areas. Blob counter calculates blobs from the greyscale image. The blob with greater size is hand region.

B. *Feature Extrction of Static and Dynamic Gestures*

After the extraction of blob from the hand area detection, blob is processed through different processes for detecting static and dynamic gestures as shown in Fig. 3.

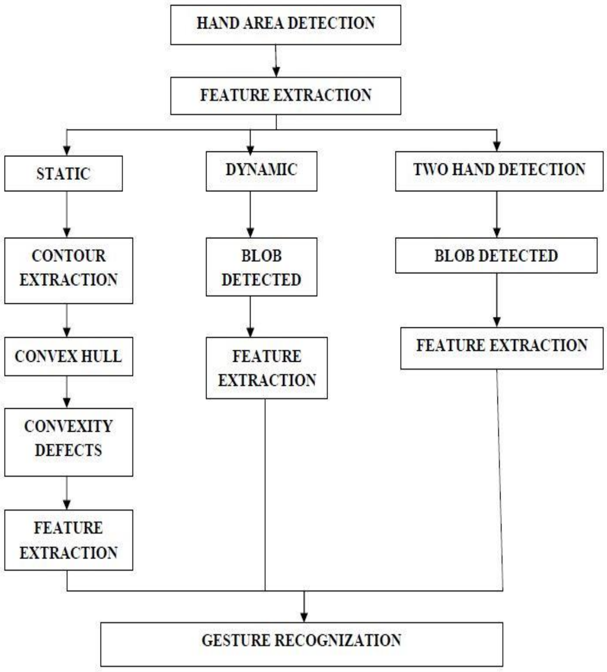

Fig. 3. Feature extraction of static and dynamic gestures.

1) *Static Hand Gesture*

For static hand gesture recognition, extracted blob is passed through different processes like contour extraction, convex hull and convexity defects. All these processes are mainly used to extract the features like subtract the background, checking whether the detected blob is hand and number of fingers detected. The features thus extracted are passed to gesture recognition module for recognition of gestures.

a) *Contour Extraction*

Contour detection is the foremost technique used for identifying edges and blobs. The contour is the locus along the boundary having a similar property like color. Contour is drawn along the border of blob as shown in Fig. 4.

Fig. 4. Contour extraction.

b) Convex Hull

After the development of contour, convex hull is drawn all over the contour. Convex hull is the collection of points that connects the contour. The convex hull covers the complete contour of the palm as shown in Fig. 5.

Fig. 5. Convex hull.

c) Convexity defects

A sequence of contour points will be fitted inside the hull when a convex hull is properly sketched across the hand. Defect points are found which is shown in the Fig. 6. That is convexity defect is a vacancy in an object sliced out from an image, which means this portion does not belong to the palm area even though it is located inside its outer boundary. From convexity defects we detect the number of fingers in the frame. Number of fingers in the frame is calculated by adding one to the convexity defects obtained from the frame.

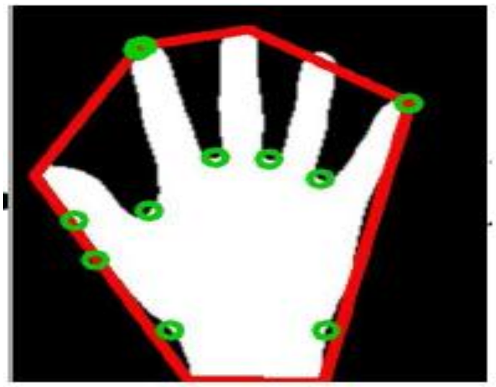

Fig. 6. Convexity defects.

2) Dynamic hand Gestures

In dynamic hand gesture, the blob extracted is processed for the dynamic hand gesture. Two dynamic gestures which has been implemented are; the movement direction of the hand and the shape detection. The movement of hand gesture is obtained by setting the frame with X-axis and Y-axis. Then the hand movement is traced along the X-axis and Y-axis to detect the direction. The shape detection was implemented by setting the highest point in the blob as (Xc, Yc). When hand moves, these points are stored in an array and this array is passed to the shape checker. Shaper checker is a function to check the shape which is drawn in the frame.

3) Two Hand Detection

In two hand detection, two largest blobs are extracted from the frame for further process; which includes contour extraction, convex hull and convexity defect for finding hand localization and number of fingers.

C. Gesture Recognition

The feature extraction method gives in-depth details regarding the finger count and shape of hand . These data are used for the identification of static and dynamic hand gestures. Different step involved in gesture is shown in Fig. 7.

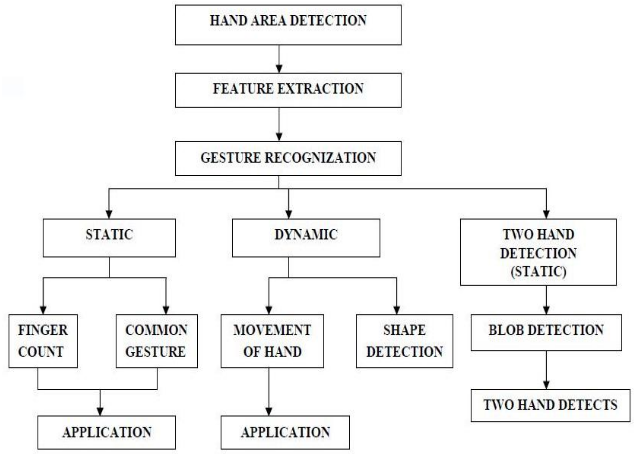

Fig. 7. Gesture recognition.

1) Static Hand Gestures

The above mentioned static hand gesture processes provides in-depth details regarding the finger count and shape of hand, using this data static hand gestures are detected. Two type of static gestures has been implemented, i.e. the finger count and common gestures.

a) Finger Count

Finger count is calculated by adding one to convexity defects. A snapshot of finger count gesture is shown in Fig. 8. i.e. Finger Count = Convexity Defects + 1

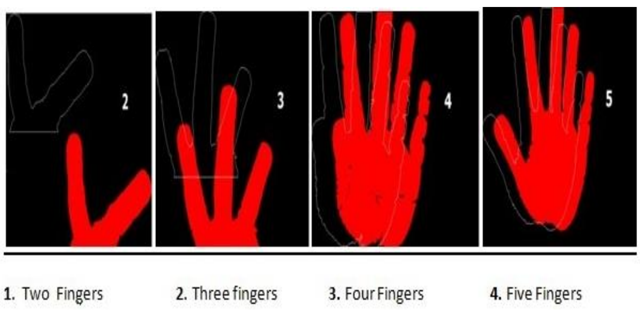

Fig. 8. Finger count.

b) Common Gestures

Common gestures are calculated based on two factors; finger count and aspect ratio.

Aspect ratio is the ratio between height and width.

i.e. Aspect Ratio = Height / Width

Four common gestures recognized in this recognition system are ‘Victory’, ‘Letter Y’, ‘Complete Hand’ and ‘I Love You’ which are shown in Fig. 9. The gestures ‘Victory’ and ‘Letter Y’ finger count is two, but their aspect ratios are different. For ‘Victory’, aspect ratio is greater than 1 and maximum of Y-convexity defect is less than 60. Whereas for

‘Letter Y’, aspect ratio is less than 0.8 and maximum of Y-convexity defect is greater than 60.

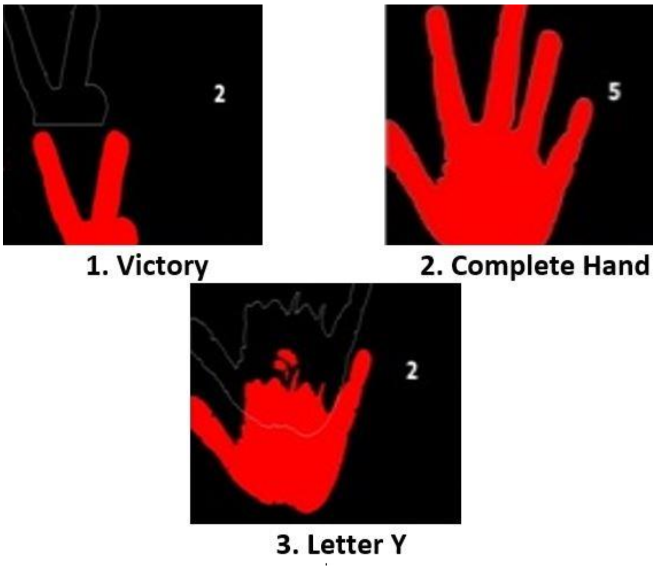

Fig. 9. Common gestures.

Similarly, for ‘I Love You’ gesture the finger count is three and aspect ratio is between 0.9 and 1.1. For ‘Complete Hand’ gesture finger count is five.

2) Dynamic Hand Gestures

The movement of blob is analyzed, which is used for extracting the dynamic gestures [6] [7]. Two types of dynamic gestures have been implemented; i.e. direction of motion hand and shape checking.

a) Movement Direction

Four directions i.e. left, right, up and down are implemented in dynamic gesture. For implementing it the whole frame is considered with respect to X-axis and Y-axis. Xstart, Xend, Ystart, Yend, Box and Boy are the variables used; Xstart and Ystart are the X-coordinate and Y-coordinate of the position from which the hand starts to move respectively. Similarly, Xend and Yend are the X-coordinate and Y-coordinate of the position at which the hand stops to move respectively. According to the value of box and boy detecting the direction of hand movement.

b) Shape Checking

For shape checking, first the highest point in blob is found and is set as (Xc, Yc). Then the movement of the point (Xc, Yc) is drawn with pen until it releases. The point (Xc, Yc) for each frame of the session is saved in an array; the stored (Xc, Yc) value is passed to the shaper checker function. Shaper checker is a function to check the shape which is drawn in the frame. The snapshot of shape recognition is shown in Fig. 10.

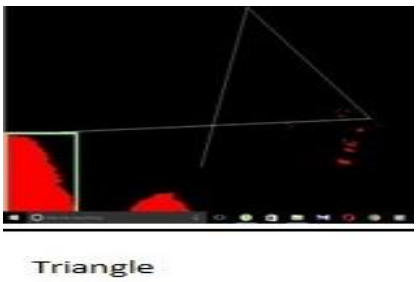

Fig. 10. Shape checking.

c) Two Hand Detection

The largest two blobs are extracted and all other processes are same as that of static hand gestures, i.e. contour extraction, convex hull and convexity defects for detecting the hand shape and finger count. Snapshot of two hand detection is illustrated in Fig. 11.

Fig. 11. Two hand detection.

d) Applications

Three applications has been made with static and dynamic gestures. With static gestures an audio player is made. Gestures like Victory, Letter Y, Complete Hand, Finger Count Three and Four are used to controls play, next, stop, previous and pause respectively. Its GUI is shown in the Fig. 12.

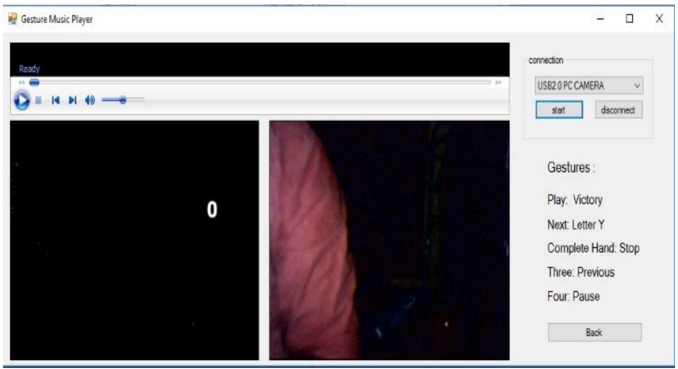

Fig. 12. Static audio player.

The dynamic gestures are used control a video player. Dynamic gestures like Left, Right, Up and Down are used to control play/ pause, stop, volume up and volume down respectively. Its GUI is shown in Fig. 13.

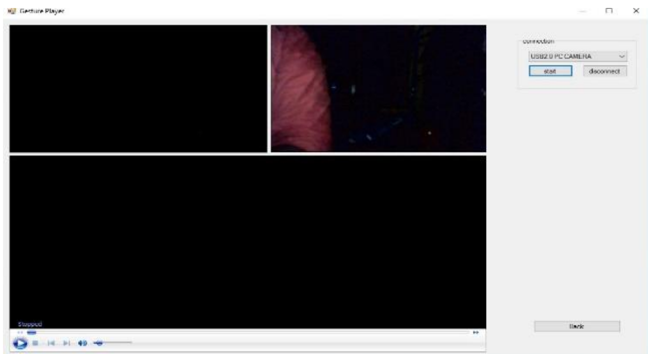

Fig. 13. Dynamic video player.

The dynamic gesture i.e. hand motion tracking is used to control mouse pointer. The highest point in blob is set as (Xc, Yc); the curser is set to the point (Xc, Yc) and map the frame to monitor size. Half the blob is click of the mouse mode i.e. when hand folds mouse pointer makes a click the object.

IV. EXPERIMENTS AND RESULTS

Testing has been done on the set of static gestures and dynamic gestures. In static gestures, finger count from two to five and other common gestures are analyzed; and in dynamic gestures direction of the hand, shape checker and two hand detection recognition are analyzed.

The test has been taken in two formats; i.e. based on the background of the frame and skin color of the hand. In testing based on the background, testing is done in three types of room based on the light; i.e. dark room, medium dark room and light room. The time for gesture recognition is analyzed for both static and dynamic gestures; the time is calculated of total rate (100%) in seconds. The time limit was set to 10 seconds i.e. 100% gesture recognition output must be obtained within 10 seconds; the results thus obtained is shown in Fig. 14 and Fig. 15.

The same is analyzed based on the skin color of hand; i.e. dark skin color and light skin color. Result is shown in Fig. 14 and Fig. 15.

Static gesture	Finger count						
	Light color hand				Dark color hand		
		Dark Room	Medium Dark	Light room	Dark Room	Medium Dark	Light room
	Rate (%)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)
Two	100	2.13	4.73	4.83	3.44	4.73	5.86
Three	100	1.74	4.24	4.46	1.97	4.24	5.56
Four	100	1.85	3.80	3.48	2.26	3.80	4.98
Five	100	1.49	3.67	3.13	3.8	3.67	4.63
Common gestures							
Victory	100	2.4	4.44	4.69	3.13	4.44	5.69
Letter Y	100	1.7	3.67	3.34	1.94	3.67	4.34
Complete Hand	100	1.5	3.56	3.56	1.85	3.56	4.2
I Love You	100	2.2	3.8	4.3	2.49	4.8	5.4

Fig. 14. Static hand gesture recognition rates and time estimates.

Dynamic gesture	Direction of hand movement						
	Light color hand				Dark color hand		
		Dark Room	Medium Dark	Light room	Dark Room	Medium Dark	Light room
	Rate (%)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)	Time (s)
Right	100	1.33	2.43	2.73	3.44	3.73	4.36
Left	100	1.24	2.34	2.86	3.67	3.52	4.56
Up	100	1.45	2.80	3.88	4.26	4.23	3.98
Down	100	1.19	2.27	3.23	3.6	3.77	4.45
Shape checker							
Triangle	100	1.3	2.43	3.49	2.13	3.44	4.69
Quadrilateral	100	2.1	2.67	3.88	1.94	3.79	4.24

Fig. 15. Dynamic hand gesture recognition rates and time estimates.

Performance evaluation is done on the basis of light color hand and dark color hand. Performance evaluation of light color hand in static and dynamic recognition tests is shown in Fig. 16 and Fig.17; an average time(s) 4s and 3.5s is obtained in static and dynamic recognition respectively.

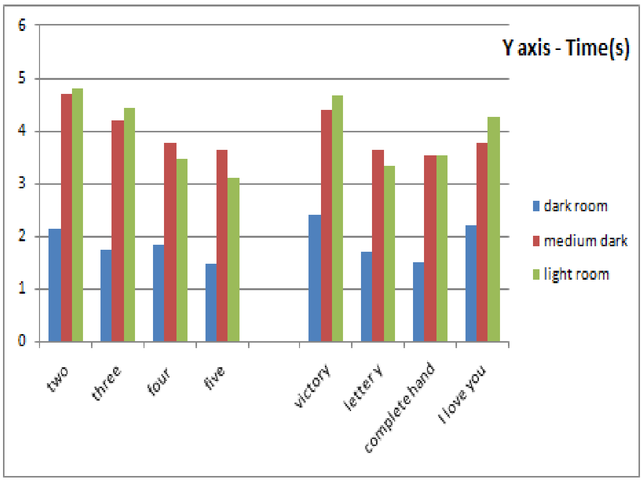

Fig. 16. Performance evaluation of light color hand in static recognition.

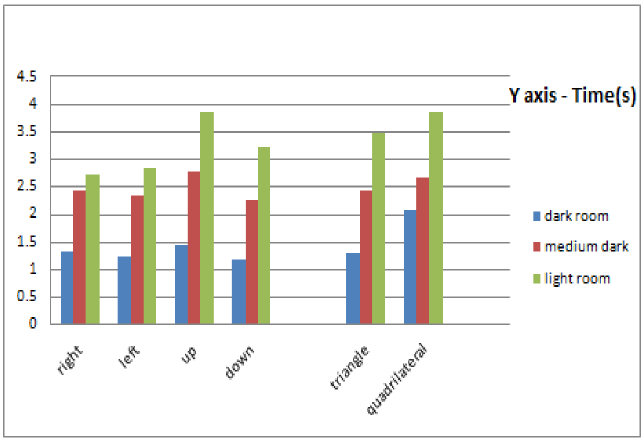

Fig. 17. Performance evaluation of light color hand in dynamic recognition.

Performance evaluation of dark color hand in static and dynamic recognition tests is shown in Fig. 18 and Fig. 19; an average time(s) 5.2s and 4.2s is obtained in static dynamic recognition respectively.

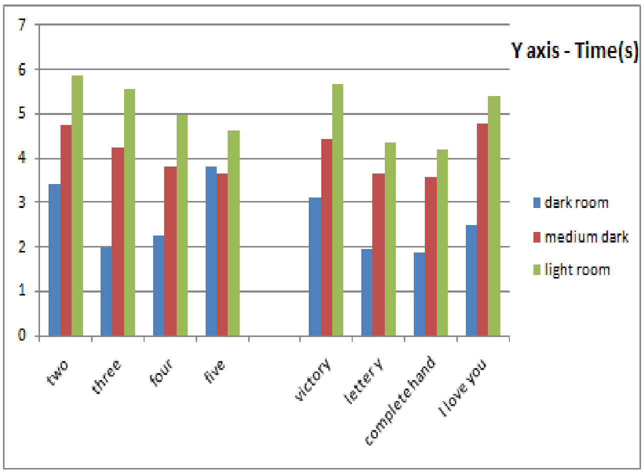

Fig. 18. Performance evaluation of dark color hand in static recognition.

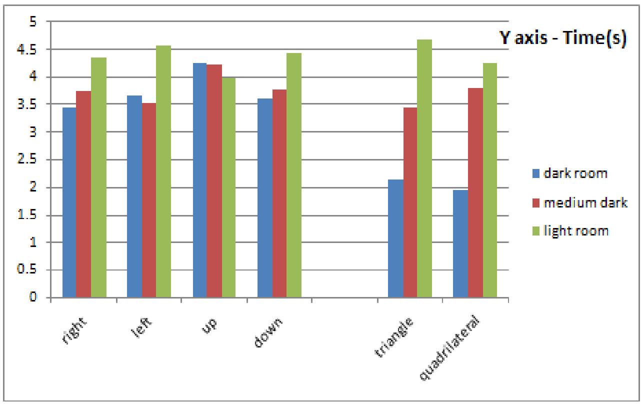

Fig. 19. Performance evaluation of dark color hand in dynamic recognition.

The results from the three scenarios i.e. light, medium and dark room conditions it can be inferred that room with less brightness and light colour hand showed more accuracy.

V. CONCLUSION

This research presents a system that is capable of interpreting both dynamic and static gestures from a end user, with the aim of implementing it in real-time human computer interaction. Thus, a hand tracking and gesture recognition system was created for HCI using cost-effective hardware. The users can interact with PC applications or games by performing hand gestures instead of depending on hardware controllers. This system can be easily calibrated to suit any skin color and background. It can trace hand and finger locations in real time and is capable of identifying simple hand gestures. The system was tested in real time situations and was able to identify all trained gestures. The experimental outcome infer that the system developed is reliable in recognizing the pre-defined commands.

REFERENCES

[1] D. K. Ghosh and S. Ari, "A static hand gesture recognition algorithm using k-mean based radial basis function neural network," 2011 8th International Conference on Information, Communications & Signal Processing, 2011, pp. 1-5, doi: 10.1109/ICICS.2011.6174264.

[2] T. Nguyen, D. Vo, H. Huynh and J. Meunier, "Geometry-based static hand gesture recognition using support vector machine," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), 2014, pp. 769-774, doi: 10.1109/ICARCV.2014.7064401.

[3] H. Duan, Q. Zhang and W. Ma, "An approach to dynamic hand gesture modeling and real-time extraction," 2011 IEEE 3rd International Conference on Communication Software and Networks, 2011, pp. 139-142, doi: 10.1109/ICCSN.2011.6014236.

[4] S. Shiravandi, M. Rahmati and F. Mahmoudi, "Hand gestures recognition using dynamic Bayesian networks," 2013 3rd Joint Conference of AI & Robotics and 5th RoboCup Iran Open International Symposium, 2013, pp. 1-6, doi: 10.1109/RIOS.2013.6595318.

[5] G. Plouffe, A. Cretu and P. Payeur, "Natural human-computer interaction using static and dynamic hand gestures," 2015 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE), 2015, pp. 1-6, doi: 10.1109/HAVE.2015.7359473.

[6] A Haria, A Subramanian, N Asokkumar, S Poddar, J S Nayak, "Hand gesture recognition for human computer interaction", Procedia Computer Science, Volume 115, 2017, Pages 367-374, ISSN 1877-0509.

[7] D S Korake, "Real Time Static and Dynamic Hand Gesture Recognition System with SURF Algorithm," International Journal of Innovative Research in Computer and Communication Engineering, 2017, Vol. 5, Issue 3, March 2017, ISSN 2320-9801.