A Review on Implementation of Virtual Mouse Conformation using Hand Gesture Tracking

Rohit Shakya[1], Ankur Taneja[2]

M.Tech Scholar, Department of Computer Science and Engineering, SAM College of Engineering & Technology,

Bhopal, M.P, India[1]

Assistant Professor, Department of Computer Science and Engineering, SAM College of Engineering & Technology,

Bhopal, M.P, India[2]

ABSTRACT: As per the technology grown, the communication between human and machine turned spectacular and touch-less. There are various method through which a human-computer interface can built such as – mouse interaction, touch-screen and gesture control. But now a virtual mouse pointer using hand tracking may affect the way of interfacing machines. Trending machines only follows certain distinct gestures to address the command for operating machines or computer. The purpose of this paper is to review and identify the flaws in related works. There are various methods which have been implemented such as MYO arm band, Convolutional Neural Network, Kinetic Sensors and many more. MYO band works with muscle movement and it has been observed that every gesture pertain specific muscle tenseness which is liable to identify the desired gesture. But MYO band has some limitations because of certain gestures which do not met the desirable control. Most of the systems are limited with the background that can work effectively only with plain background rather than complex one. A system is required which can work effectively with all kind of backgrounds without any interruption.

KEYWORDS: Virtual Mouse, Hand Gesture Recognition, MYO Arm Band, Convolutional Neural Network, Kinetic Sensors.

I. INTRODUCTION

It has been observed that, use of Computer in day to day life, either at the work place or at home is rapidly growing. Various technologies have been developed in order to simplify the interaction of computers and human. There are two devices which are essentially important for the communication between computers and human i.e. keyboard and mouse. Each innovation in this field is considered as an effort towards the enhancement of technologies which can enable the complex communication between computer and human. There are so many successful result oriented inventions in the field of human compute interface has made to enhance the communication between computer and human. Requirements of human which are intricate too are constantly rising which in turn required an efficient computer having capability of high level of programming and intuitive knowledge that can meet human's need. Role of programmers or coders played a vital role to easier the use of computer for human. Every new technique or product comes in the market, targeted to reduce the effort made to execute any complicated job. Gestures have an extensive range of functions so that to conclude the word gesture with a definition is not easy. Typically, Gesture is considered as signs made with the help of hands or by using some facial expressions intended to communicate with the others without any verbal conversation. To build a successful and proper communication, both transmitter and receiver should be aware of same set of data which is being used to form a gesture. Basically, forms of Gesture is classified into two various groups which are dynamic gestures and static gestures.

Fig.1: HCI [1]

Dynamic gestures are those which can be modified with time. Static gestures are considered as a sign which is permanent for some messages. Waving of hand can be taken as a gesture of saying either Goodbye! Or Hi! This is an example of dynamic gesture. Plus sign is used for ambulance which lies under static gesture. To recognize the whole information in a message which is being send using gestures needs to understand both static and dynamic form of gestures. The process used to extract the information from the visual signs is called gesture recognition by using predefined database.
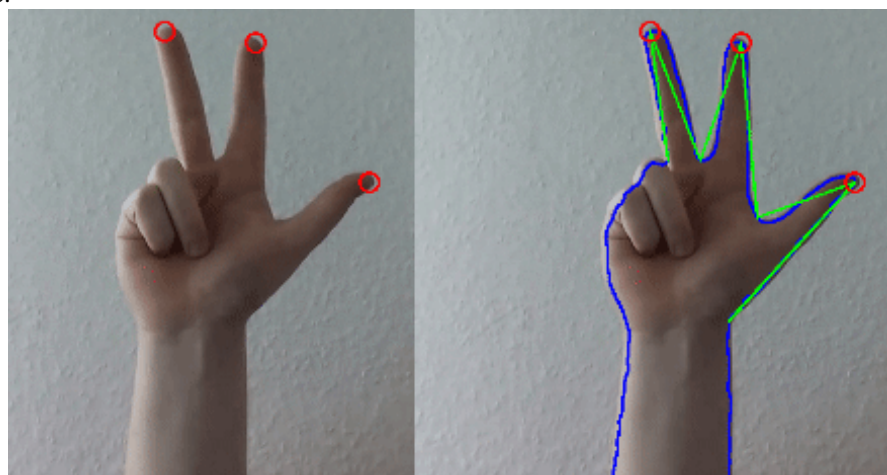

Fig. 2: Hand Gesture [2]

Virtual mouse is an application of software intended to operate the features of actual mouse in a computer without using real device of mouse. To implement the technique of virtual mouse, a normal webcam is required to examine and capture the gestures made to operate the functionality of a mouse. Generally, the function of clicking, drag-drop and right-left click is performed by using a hardware device called mouse or by using keyboard. Virtual mouse removes the requirement of the hardware device and just by using some trained and predefined gesture and image processing technique.

Fig. 3: Virtual Mouse Interaction [3]

II. RELATED WORK

Zhi-hua Chen et al. proposed a system which is based on OpenCV library by segmenting palm and fingers to recognize gesture. The structure employed by the system uses background subtraction technique to take out the region of hands only. Segmentation of palm and finger separately took place through which gesture can be recognized through finger by wrist pointing and palm masking. To identify the tags of hand gesture, rule classifier was used. But this system also requires plain background for segmentation and recognition; it does not work correctly in non-plain background as the performance of the system which has been proposed typically relied on the outcomes of hand identification. If the extracted image consisted of object having identical color as that of skin, it may interrupt the overall performance of the system as the detection technique used for hand gesture recognition became degraded. The overall computed accuracy of the system which has been proposed in this paper was 96.6% for plain background which get enhanced in the proposed system [4].
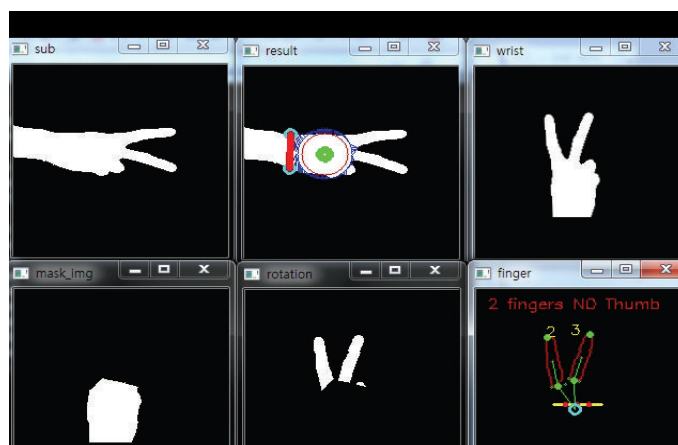

Fig. 4: Segmentation and Recognition [4]

Atharva Ajit et al. proposed a system which is based Microsoft kinetic sensor that is a physical sensor through which hand gesture is separated and later match with directory to recognize input gesture. Microsoft kinetic sensor is using infrared to input data. System calculates centroid point in palm to determine the position of palm. The system is bit costly because of the sensor and uses hand geometry to recognize gesture which may have more error rate with less accuracy. The method which has been projected in the system provided an application which can be utilized to monitor and offer surveillance concerning the old people and human – computer communication. The information acquired from the deployed kinetic sensor used to discover the viability of gesture recognition. Method used the curves and lines of palm and understands the form made by hands using Dynamic Time Warping (DTW) method. Those curves and lines of palm are extracted by employing the technique of geometric feature on palms [5].
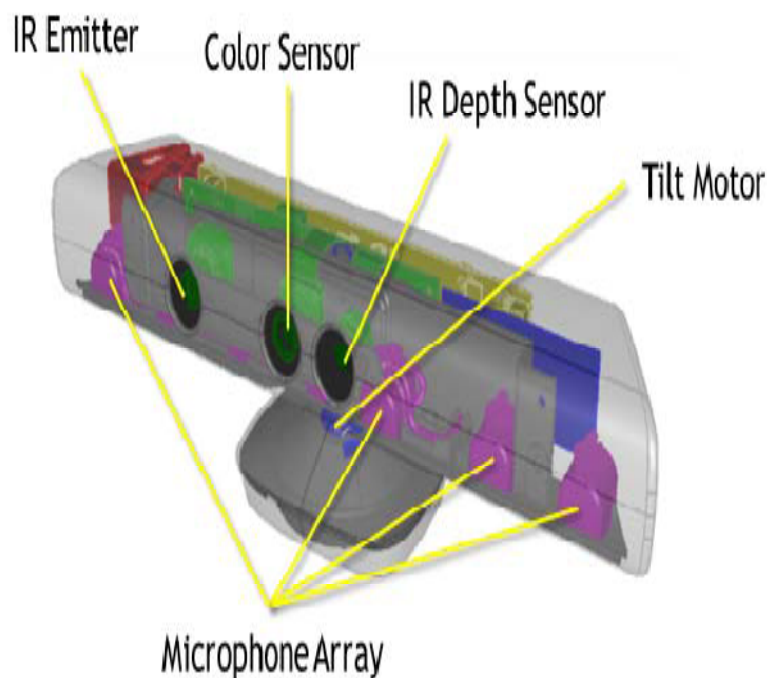

Fig. 5 Kinetic Sensor [5]

Marco et al. proposed a system which is based on MYO Armband which can analyze or recognize muscle activity and on the basis of that gesture has been recognized. MYO Armband is costly and certain gestures can recognize only. There is a muscle movement or muscle stretches while putting a particular gesture and MYO Armband can recognize that gesture and can perform action accordingly. The system which has been proposed in the system extracted the data obtained by the surface electromyography i.e. EMG computed on the muscles of the fore-arm by using MYO armband. System provided a training process to understand the process of gesture recognition to every user before using it. It is required in the system which has been proposed to record the process for 5 times with the duration of 2 seconds. To organize the EMG which is visible through a window exploited the K-nearest neighbor algorithm and the dynamic time wrapping algorithm. A detector is also induced in the system to monitor the activities performed by the muscles to modify the accuracy of the system. The accuracy obtained by the system of gesture recognition which has been proposed in this paper is 89.5%. But the general utilization of system is still costly due to the technique used in the proposed method [6].

Fig. 6: MYO Armband [6]

Pei Xu et al. has introduced a system in the paper which is also based on convolutional neural network but along with Kalman filter which is able to estimate hand position. Here, system is controlling mouse events as virtual mouse by using hand gesture. This system separates palm and finger to recognize finger based recognition. There are three various factors on which the overall functionality of the developed system relied i.e. detection of hand, recognition of gesture and interaction between human and computer. The developed technique took vigorous control to operate the functionality of mouse and keyboard with improved accuracy through gesture recognition. By employing monocular camera gesture made by hands was captured and identified by exploiting the technique of CNN i.e. convolutional neural network. It has been stated in the paper that intricate gesture can be identified by the system. Kalman Filter is also exploited in the technique which has been proposed in the paper to compute the positions of hand which is used to control the movement of cursor of mouse [7].
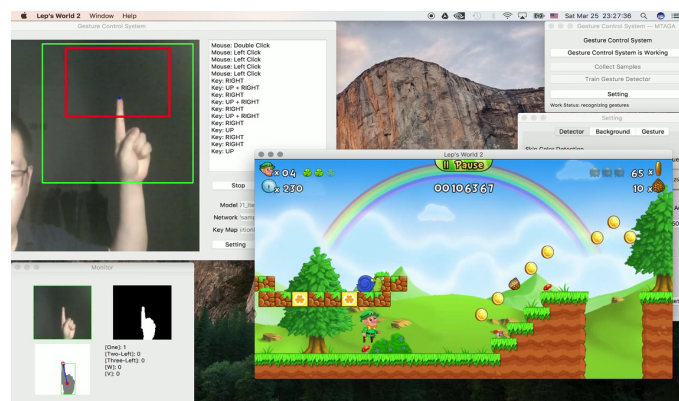

Fig. 7: HCI through Plain Background [7]

Aashni Hariaa et al. introduced a system in the paper which relies on contour extraction for creating virtual mouse through which particular application can be controlled. This system intended to develop a vigorous hand gesture identification technique which has the capability to identify the gesture made by hands whether it can be static or dynamic. System is trained to perform action based on different gestures made through hands. After extraction of the

gesture made by hands, system is trained to open websites or applications like VLC Player, Google Chrome and MS power point. Training of dynamic gesture is done in order to swap the slides shown in power point presentation. In the gesture recognition system which has been developed in the paper, trained seven different gestures. Among those seven gestures, six are static gesture and one gesture is dynamic [8].
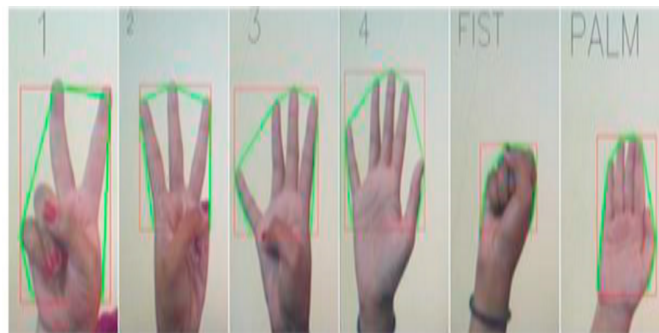

Figure 8: Static Gestures Used In Gesture Recognition System [8]

Since the first gesture having "V" symbol is trained to open VLC player. Second image shows the sign of two which is used to start default browser having homepage of Google. Third Sign is used to open YouTube. Fourth indication which shows five fingers is used to close those applications which are opened in front. Fifth sign in the figure shows a fist which will launch the Microsoft power point application. The last static figure has a sign of normal hand showing palm which is used to toggle the network of Wi-Fi. In this paper developed system stated that, the requirement of hardware for the execution of system is reduced in the technique. But the problem with this system is reduced accuracy rate in complex background. It can only work correctly with plain background. Extraction of only contour is not an effective way to develop a reliable controlling system. In figure which is shown below, suitable contours of various gestures used in the system which was presented in this paper. Since, contour has been exploited in any recognition system to draw the edges of the extracted image [8].

Peijun Bao et al. developed a system which is based on convolutional neural network where hand segmentation is not prefered, instead of that a network has been trained which is able to recognize hand gesture. Method which has been proposed in the paper exploits the convolutional neural networks to straightly categorize the gestures made by hands in the extracted image. System didn't use any typical sets of process for recognition like process of segmentation, preprocessing which have the feature to dispose any unrelated area which do not contain any portion of hand. Since the intended system which contains the network of recognition process, categorized seven different forms of gestures made by hands which is not dependent not the user. System didn't induce any framework selection of the extracted image to recognize the formed gestures. Method which has been proposed exploited deep CNN process for the recognition purpose. It has been stated in the paper, that the proposed system can recognize the gesture in any background. But it can be observed in the image mentioned below, that frames formed by the proposed system and the actual hands are at different places. So, the outcome may fluctuate the overall result of the system. The technique used in this system is not very effective in complex background; it works well with plain background. As shown in the figure, that frame in which man wore red t-shirt shows a gesture but system didn't identified that gesture and marks the box at wrong place.Since the use of deep CNN i.e. convolutional neural network process to denoise the image and to recognize the hand gesture is trained for limited sets of gesture and can able to handle limited noise. So, the utilization of system which has been proposed in real time will not be a reliable option [9].
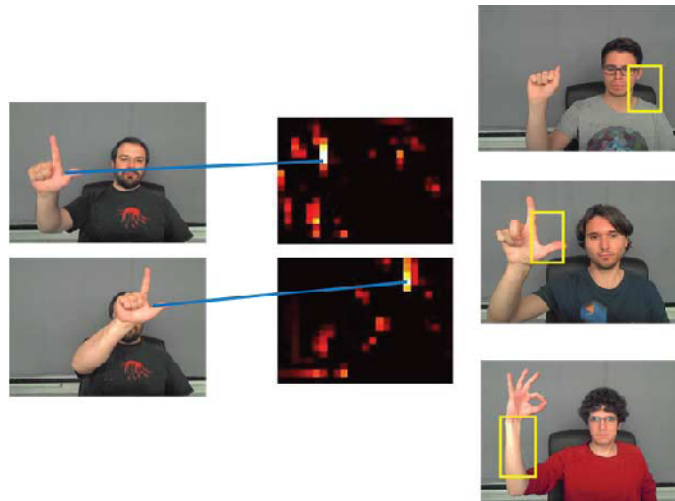
Fig. 9 CNN based system [9]

III. PROBLEM IDENTIFICATION

Most of the existing systems are based on MYO Armbands in which system is trained according to the activity of muscles when any kind of gesture is made through hands. Since every gesture has different activity of muscles so the system limits the form of gestures made through hands for recognition. As system needs to get trained about the activity of muscles to analyze the formed gesture, it will restrict their inquiry. The technique used, also increases the cost of the system. Few systems are based on the convolution neural network where the detection of hand gestures are very accurate but to maintain the accuracy, used background needs to be plain as system got confused between the non plain background and gestures. In those systems, it becomes problematic to correctly point out the positions and gesture in complex background. Some of the systems are based on contour extraction where the same problem suffers that is background complexity. It is required to develop a system that can work in complex background with high level of accuracy, through which interaction with computer will become easy and effective [1-7].

Fig.10: HCI in plain background [9]

IV. CONCLUSION & FUTURE SCOPE

Thus the survey of all these systems concluded at a point that hand gesture recognition does not work well with complex background or non-plain background. All the systems are lacking somewhere due to higher error rate in non-plain background. A system requires that can fulfill these bugs and a recognition would be possible with complex background with high level of accuracy. The current proposed concept of hand gesture recognition can be enhanced in future by developing a system that will be able to recognize hand gesture from complex or plain background with high level of accuracy similar with plain background.

REFERENCES

[1] https://www.pantechsolutions.net/blog/top-10-gesture-recognition-projects-with-free-source-code/
[2] https://medium.com/@muehler.v/simple-hand-gesture-recognition-using-opencv-and-javascript-eb3d6ced28a0
[3] J.L. Raheja, M. Minhas, D. Prashanth et al., "Robust gesture recognition using Kinect: A comparison between DTW and HMM", Science Direct, 2015
[4] 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.
[5] D. H. Pal and S. M. Kakade, "Dynamic hand gesture recognition using kinect sensor," 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), Jalgaon, 2016, pp. 448-453.
[6] M. E. Benalcázar *et al.*, "Real-time hand gesture recognition using the Myo armband and muscle activity detection," 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), Salinas, 2017, pp. 1-6.
[7] Pei Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System", IEEE Transaction, 2017.
[8] Aashni Hariaa, Archanasri Subramaniana, Nivedhitha Asokkumara, Shristi Poddara , Jyothi S Nayaka, "Hand Gesture Recognition for Human Computer Interaction", Science Direct, 2017.
[9] P. Bao, A. I. Maqueda, C. R. del-Blanco and N. García, "Tiny hand gesture recognition without localization via a deep convolutional network," in IEEE Transactions on Consumer Electronics, vol. 63, no. 3, pp. 251-257, August 2017.