Intangible Hand Gesture Based Human-Computer Interaction System

Anjali Singh[1], K. Ezhil[2], K. Venkatesh Guru[3]

[1, 2] Student, Department of Computer Science & Engineering, K.S.R. College of Engineering, Tamil Nadu, India.
Email: rajpootanjali1620@gmail.com, zhil047@gmail.com

[3]Assistant Professor, Department of Computer Science & Engineering, K.S.R. College of Engineering, Tiruchengode 637215, Tamil Nadu, India. Email: guru2ksr@gmail.com

Abstract - Human-Computer Interaction (HCI) may be a multidisciplinary field of study specializing in the look of engineering and, particularly, the interaction between humans and computers. The computer technology is growing continuously, the need for natural communication between humans and machines also increases. The communication between the user and the computer can be established through various input devices such as the keyboard, mouse etc. Input devices such as mouse is very useful for device control, this could be inconvenient for people who are not used to it for interaction. The proposed method in this paper uses a webcam which captures the gestures provided as input by the user, input is processed further, and functions related to that gesture is carried. Using OpenCV, the captured video is crushed down into endless image frames. The captured images are processed, and the gesture is detected. Using the OpenCV library, the cursor movement by hand gesture is done, uses Python programming language, which maintains an ease to grasp code through its primitiveness. Python modules such as PyAutoGUI and packages such as NumPy are used here. Various mouse operations like cursor movements, right click, left click, speed of the cursor, drag and drop have been performed. We have effectively tried our system for an intangible interface between human hand and PC with less complexity.

Keywords - Hand gesture, OpenCV, HCI, Convexity defects, PyAutoGUI module

I. INTRODUCTION

The fundamental expectation of Human Computer Interaction is to improve the intercommunication among clients and PCs by making the PC increasingly acceptant to client needs. Today Human-Computer Interaction isn't just delineated to console and mouse communication. Diverse tangible modes, for example, motion, discourse, body and outward appearance characterizes the communication between people. Having the option to collaborate with the framework normally is getting always significant in numerous fields of Human Computer Interaction. Correspondence among human and machine stays a troublesome errand. Motions for prevailing gadgets might be an innovation by and by to be abused. Our hand is utilized wide as a characteristic human to human connection by signaling. It might be misused for the correspondence among human and machine moreover. There is a fast improvement in Human-Computer Interaction, the manner in which we connect with figuring devices. We are adjusting to the continually expanding requests of current ideal models. One of the most helpful instruments in such manner is the absorption of Human-to-Human motions which encourage correspondence and declaring thoughts. Motion acknowledgment requires joining of signals, stances, and developments for imparting or passing on specific messages.

Because of the way that the human hand stances and signals are as yet an incredible entomb human correspondence methodology the examination network keeps an enthusiasm in regard to the utilization of the hand motion as a way to control or to impart through it with fake frameworks. A perfect info interface ought to perform very much dependent on some normal criteria (i.e., precision, throughput, reasonableness, amiability, portability, ease of use). In light of these criteria, human motion has gotten mainstream as HCI interface and its utilization is expanding step by step. Hand Gesture Recognition has extraordinary pertinence for HCI inferable from a rundown of utilizations in different spaces, which incorporates communication via gestures acknowledgment, augmented reality, and PC games. Gesture based communication is the widely utilized strategy when sound transmission is restricted or very improbable. Hand signs are additionally every now and again utilized when techniques for composing and composing are troublesome, however the mode of vision is uncluttered.

There are five segments in this paper. Segment 1 contains the related works; Section 2 is enveloped with the proposed framework; Section 3 has the test results; segment 4 contains the usage and results; Section 5 at last has the end.

II. RELATED WORKS

Some relevant studies are discussed in this section to gather more information and also implement it from the same.

A hand-mouse interface that introduces a new concept called "virtual monitor" has designed by Changhyun Jeon [1], to extract a user's physical features through Kinect in real-time. This virtual monitor allows a virtual space to be controlled by a hand mouse. It is possible to map the coordinates on the virtual monitor with the coordinates on the real monitor accurately. The outstanding intuitiveness maintained by the hand-mouse interface that was based on the virtual monitor concept, which is the strength of the previous study and enhances the accuracy of mouse functions. To evaluate the intuitiveness and accuracy of the interface, we experimented with 60 volunteers ranging from teenagers to those in their 50s. The results of this experiment showed that 84% of the subjects learned how to use the mouse within 1 min. Besides, the experiment showed the high accuracy level of the mouse functions [drag (80.8%), click (80%), double-click(76.8%)]

Feng-Sheng Chen [5] proposed a hand motion acknowledgment framework to perceive constant signals before the stationary foundation. The framework comprises of 4 modules: ongoing hand following and extraction, highlight extraction, Hidden Markov Model preparing, and motion acknowledgment. To begin with, we applied a continuous hand following and extraction calculation which follows the moving hand and concentrates the hand area, at that point we utilize the Fourier descriptor which describes spatial highlights and the movement investigation for portraying the fleeting highlights. We joined the spatial and world highlights of the info picture arrangement as the element vector. In the wake of separating the element vectors, we apply HMM to perceive the information signal. The signal to be perceived is scored independently against various HMMs. The model which has the most elevated score shows the relating motion. In the examinations, we have tried our framework to perceive 20 distinct motions, and the perceiving rate is above 92%.

M. F. Shiratuddin [6] proposed the Multi-Hand Gesture (MHG) cooperation worldview, a non-contact motion acknowledgment framework is utilized to identify progressively the hands and fingers developments, and furthermore their situations in 3D space. These signals are then deciphered and used to execute explicit directions or assignments. The paper by M. F. Shiratuddin proposes a structure for a non-contact Multi- Hand Gesture collaboration strategy for compositional plan, and it likewise talks about 1) a non-contact MHG connection procedure and the WIMP, and 2) the improvement of an early model of a non-contact MHG acknowledgment frameworks. The model framework comprises of programming which is created to perceive and decipher the multi-hand motions caught by the Kinect sensor which does constant filtering of three-dimensional profundity and space.

The paper by R. Lulu1[7] manages the improvement of human cooperation with the advanced world. The scaling down of processing gadgets permits us to be in constant touch with the computerized world. Limitation of data on conventional stages like paper, computerized screen, and so on is being overwhelmed with the assistance of this given innovation. Reliance on customary equipment input gadgets like mouse, console, and so on will be decreased altogether, consequently permitting convey ability. It utilizes hand developments or signals to give contribution to a PC or some other electronic or computerized gadget. The Virtual Keyboard is an adjustment of the sixth sense innovation created by Pranav Mistry, a Ph.D. understudy at MIT media labs. It chips away at an undifferentiated from rule of hand development or the motion control to perform basic capacities like Right snap, Left snap, Scroll control, and so on, which would some way or another must be performed with the assistance of equipment gadgets like mouse or console. The picture preparing strategy is a significant reason for the execution of this innovation. The key highlights of this innovation incorporate various controls like media player volume control, PowerPoint slide control, camera control, looking over mouse control, commencement and end of a call. The paper manages the computerized world registering gadgets with the improvement of human association which permits us to be in constant touch with the advanced world. This code will be created in MATLAB.

The paper by Dulayatrakul J [13], A strong hand signal acknowledgment calculation for remote human-machine connection proposes which has been streamlined for usage on an implanted stage. The normal issues identified with varieties in lighting conditions defeat by Hue-immersion esteem (HSV) thresholding and unit-slope vector (UGV) foundation subtraction strategies that are utilized. The fingers and hand signals identified by Top-cap change, which are meant order contributions for remotely controlling any media player. The calculation performs proficiently and exactly on an implanted board with a normal computational expense of 154 milliseconds for every signal and is vigorous to changes in delineation that shows in Experimental outcomes.

This paper utilizes 'A Sliding Window Approach to Natural Hand Gesture Recognition utilizing a Custom Data Glove' investigates the acknowledgment of hand motions dependent on an information glove outfitted with movement, twisting and weight sensors. Granit Luzhnica [11] and group chosen 31 normal and collaboration arranged hand signals that can be obtained for universally useful control and correspondence with figuring frameworks. This information glove is exceptionally assembled and contains 13 twist sensors, 7 movement sensors, 6 weight sensors, and a magnetometer. We present the information assortment try, just as the plan, assessment, and choice of an order calculation. As we utilize the sliding window way to deal with information handling, our calculation is appropriate for stream information preparing. Calculation choice and highlight designing brought about a mix of straight discriminant investigation and strategic relapse with which we accomplish an exactness of over 98.5% on a persistent information

stream situation. While expelling the computationally costly FFT-based highlights, we despite everything accomplished an exactness of98.4%.

In this paper, Lee DH[17], proposed a continuous hand motion acknowledgment framework dependent on the picture entropy acquired utilizing a stereo camera. In existing frameworks, hand recognition has been basically led in a specific obliged condition. Be that as it may, in this framework, they have actualized an acknowledgment framework for approaching hand pictures progressively. In the perception step, we have applied a profundity map utilizing the SAD strategy dependent on right-left pictures procured utilizing a stereo camera. The frontal area article and hand identification execution saw in this framework. The distinction picture entropy of the info picture and the normal picture is utilized by hand acknowledgment framework. We have executed an acknowledgment explore utilizing a hand motion database (size 240) to assess the presentation of proposed innovation. The proposed strategy has a normal acknowledgment pace of 86% that appeared in the trial results.

III. PROPOSED SYSTEM

An epic immaterial hand signal based PC mouse control framework, utilizes a webcam through which motions gave by the client are caught, prepared and the capacity identified with that motion is completed. The paper expects to recognize a hand signal and play out the relating cursor activity. A three-organize module is utilized to actualize it. To start with, the movement is recognized. At that point the edges are identified and the hand shapes are determined. Later the signal is named with the comparing cursor activity.

The proposed framework comprises of 3 significant modules,

1) Motion identification
2) Feature extraction
3) Gesture acknowledgment

The structure, conduct, and perspectives on a specific framework is characterized by a framework engineering which is a reasonable model. A design depiction is the proper portrayal and portrayal of the framework, composed such that supports thinking about the structures and practices of the framework. Figure 1 shows the framework design of the hand motion based human-PC communication framework. The motion performed is caught by the webcam and is changed over into video outlines. The video outlines experience a progression of preparing and are prepared for playing out's the necessary hand signal.

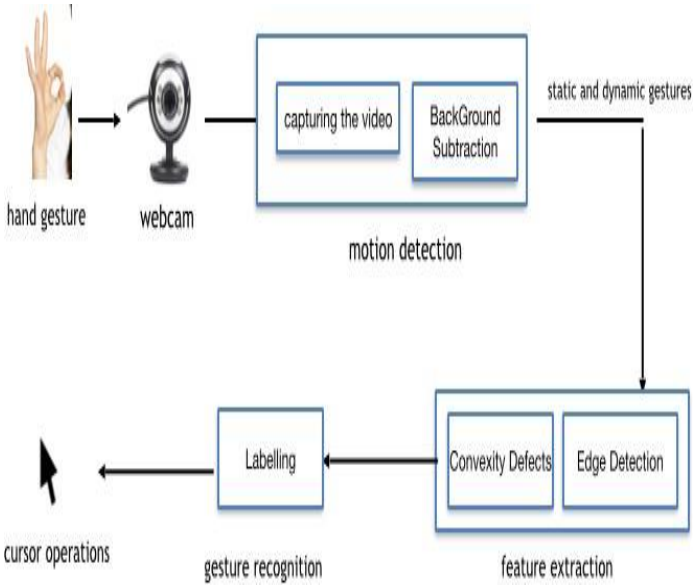


Fig. 1 System Architecture

A Functional Architecture is a model that recognizes framework work and their collaborations. It characterizes how capacities will work together to play out the framework crucial. Figure 2 shows the utilitarian engineering chart of hand motion based

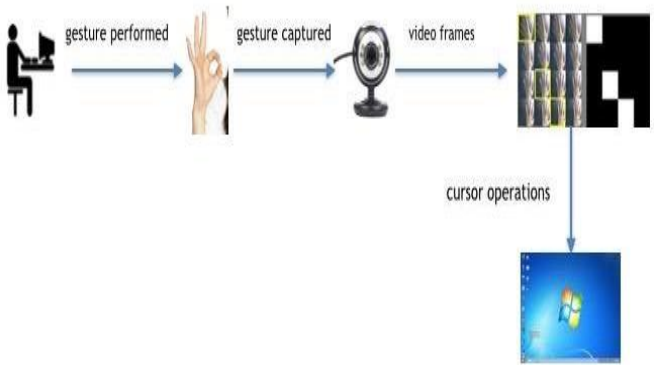


Fig. 2 Functional architecture for labeling

Human-computer interaction system. We can explain our paper with the help of two functional architecture. Figure 2 shows the architecture for labeling the required hand gesture with the corresponding action and Figure 3 shows the architecture for performing the required cursor operation.

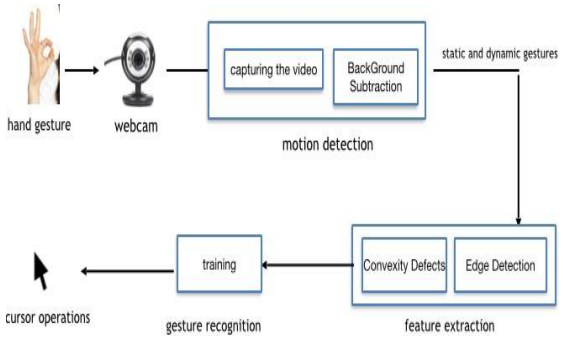


Fig. 3 Functional architecture for training

A. Camera module

This module is responsible for connecting and capturing input through the different types of image detectors and sends this image to the detection module for processing in the form of frames. The commonly used methods of capturing input are data gloves, hand belts, and cameras. In our system, we use the webcam inbuilt which is cost-efficient to recognize both static and dynamic gestures. The system has a suitable provision to allow input from a USB based webcam as well. The image frames obtained are in the form of a video. The video is subjected to different image processing techniques such as color conversion, noise removal, thresholding following which the image undergoes contour extraction.

Noise removal: The input image, which is in RGB color space, is cropped to a size of 300 * 300 pixels. It is then converted into a grayscale image. Noise in images can be defined as a random variation of brightness or color information that is usually produced during the image acquisition process from the webcam. This noise is an undesirable aspect of the image and needs to be removed. To do this, the Gaussian filter is applied. Gaussian filtering is performed by the convolution of the Gaussian kernel with each point in the input array. These are then added to produce the output array.

Thresholding: Thresholding, which is a simple segmentation method, is then carried out. Thresholding is applied to obtain a binary image from the grayscale image. The thresholding technique compares each pixel intensity value (I) concerning the threshold value (T). If I<T, the particular pixel is replaced with a black pixel and if I>T, it is replaced with a white pixel. A threshold value (T) of 127 is used in our work which classifies the pixel intensities in the grayscale image. The maximum value of 255 is the pixel value used if any given pixel in the image passes the threshold value. The two types of thresholding that are implemented are Inverted Binary Thresholding and Otsu's Thresholding. Figure 4 shows the threshold images of the hand gesture.

Fig. 4 Threshold image

B. Feature extraction

Contours are useful for object detection and recognition in image processing. In our work, we have used contours, to detect and recognize the hand from the background. The curves that link continuous points, which are of the same color, are called contours. The second step is to draw the contours which can be used to draw any shape provided the boundary points are known. Later, the convex hull is found. Mathematically, the convex hull of a set X of points in an affine space is defined as the smallest convex set that contains X. Any deviation of the object from this convex hull can be considered as a convexity defect. The convex hull is found by identifying the fingertips. If the image contains defects, then convexity defects are found according to which the gesture is detected. If there are no defects, then the image is classified using Haar cascade to detect the gesture. For gestures like palm and fist where there are no convexity defects, the Haar cascade classifier is used. Haar training will utilize a minimum of 100 images of size 20 * 20. The generated XML file is used as a cascade classifier to detect objects in OpenCV.

Applying Cosine Rule to find an angle for all defects between fingers is given by angle =

math.acos((b*2 + c2 - a*2)/(2*b*c)) * 57

To find the length of all sides of the triangle

a = math.sqrt((end[0] - start[0])*2 + (end[1] - start[1])*2)

b = math.sqrt((far[0] - start[0])*2 + (far[1] - start[1])*2)

c = math.sqrt((end[0] - far[0])*2 + (end[1] - far[1])*2)

Some gestures in our recognition system with their appropriate contours and convex hull are shown in Figure 5.

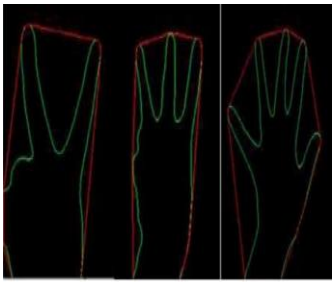


Fig. 5 Contours

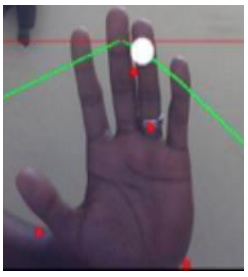


Fig. 6 Gesture recognition

This module is answerable for mapping the identified hand signals to their related activities. These activities are then used to perform cursor activity. The goal of the help vector machine calculation is to discover a hyperplane in N-dimensional space (N—the number of highlights) that particularly characterizes the information focuses. Hyperplanes are choice limits that help characterize the information focuses. Information focuses falling on either side of the hyperplane can be credited to various classes. Additionally, the component of the hyperplane relies on the quantity of highlights. SVM calculation is along these lines used to group the various signals to play out a cursor activity. Utilizing the SVM calculation, the motions are arranged and marked with their relating activity.

The yield from the component extraction module is the directions' qualities. Utilizing this classifier calculation, we arrange these facilitate values and for each ordered point, the ideal signal is mapped. We can prepare the framework with more information sources. Here, we have planned the framework to play out a right-click, left-snap, drag, and

drop, double tap. Following are the gestures,

Five finger	-Cursor movement
Three-finger	-Left click
Four finger	-Right-click Three-
finger +Five finger	-Drag and drop
Three-finger, twice	-Double click

IV. IMPLEMENTATION ANDRESULTS

In our gesture recognition system, we have included a total of three gestures, where two of them are static gestures and one is a dynamic gesture. These static gestures are shown in diagram Figures 7 and 8. The system was tested indoor (classroom and reading room) with good lighting conditions during day and night. The distance of the hand was about 1-2 feet from the screen/webcam, for which we got about 97% accuracy as we have tested the system with different people and with several trials (ex. For testing 10 times, 9 trials were accurate). At 5-6 feet distance, its accuracy was about 60%. This limitation we expect to improve in the future so that gestures can be recognized from more practical scenarios. We tested with different people. In the figure, the mouse cursor movement is shown.

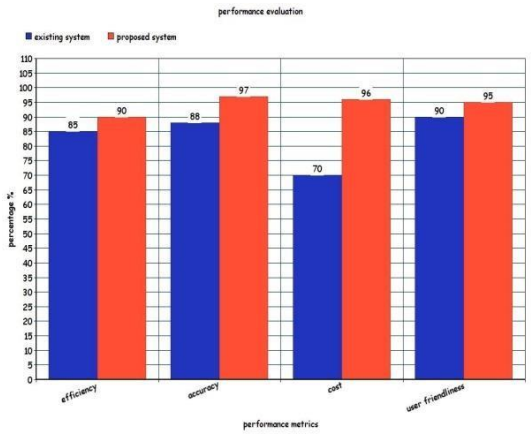
Fig. 7 Left click operation

Fig. 8 Right-click operation

V. PERFORMANCEANALYSIS

Performance analysis is a measure of the success or failure of a project using various parameters. It helps in developing a positive culture of project management that yields excellent results. A good program performance typically needs proper management of stakeholders.

VI. CONCLUSION



We have presented a method for detecting hand gestures based on computer-vision techniques, together with an implementation that works in real-time on an ordinary webcam. The method combines skin-color filtering, edge detection, convex-hull computation, and rule-based reasoning with the depths of the convexity defects. We had reported as well as user experiments on the detection accuracy of the developed prototype, detecting correctly nine in ten hand gestures made on either hand, in a self-restrained environment. We were able to create a robust gesture recognition system, hence making it more user-friendly and low cost. In this gesture recognition system, we have aimed to provide gestures, covering almost all aspects of HCI such as system functionalities, launching of applications and opening some popular websites. In the future, we would like to improve accuracy further and add more gestures to implement more functions. Finally, we target to extend our domain scenarios and apply our tracking mechanism into a variety of hardware including digital TV and mobile devices. We also aim to extend this mechanism to a range of users including disabled users.

VII. FUTURE WORK

As future work, we plan to add in the gesture detection phase an estimate of the width of each finger. This allows us to determine whether a single finger is elevated at that position or whether multiple fingers are elevated but held together. The finger-width can be calibrated for each person by measuring the width of the hand base itself and assuming that anything that has the width between one-sixth and one-fourth of the base width is a single finger. The number of fingers in a wider block can be estimated as the width of the block (computable from the points used for finger counting at present) divided by one-fifth of the base width, rounded down to the preceding integer value. Our hand gesture recognition system is not limited to the only mouse. It can be used for controlling many other devices, such as TV, robots inside a hazardous nuclear reactor in a convenient way for humans and also in some other industrial setup. It can further be applied to control multimedia applications, multiplayer gaming, etc. The sensibility of the system to lighting and background conditions can also be improved.

REFERENCES

[1] Changhyun Jeon, Oh-Jin Kwon, Dongil Shin, and Dongkyoo Shin, December 5, 2017, Hand-Mouse Interface Using Virtual Monitor Concept for Natural Interaction', IEEE access, vol. 5, ACCESS.2017, 276840525181–25188, 2017.

[2] S. Kang, C. Kim, and W. Son, ''Developing user-friendly hand mouse interface via gesture recognition,'' in Proc. Conf. Korean Soc. Broadcast Eng., pp.129–132, 2009.

[3] A. A. Argyros and M. I. A. Lourakis, ''Vision-based interpretation of hand gestures for remote control of a computer mouse,'' in Proc. ECCV, pp.40–51, 2006.

[4] W. Bland, T. Naughton, G. Vallee, and S. L. Scott, ''Design and implementation of a menu based OSCAR command-line interface,'' in Proc.21st Int. Symp. High Perform. Comput. Syst. Appl., Washington, DC, USA, p. 25, May 2007.

[5] Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang, "Hand gesture recognition using a real-time tracking method and hidden Markov models", 2003 Elsevier Science, 20 March 2003, DOI:10.1016/S0262-8856(03)00070-2.

[6] M. F. Shiratuddin and K. W. Wong, ''Non-contact multi-hand gestures interaction techniques for architectural design in a virtual environment,'' in Proc. Int. Conf. Inf. Technol. Multimedia, pp.1–6, 2011.

[7] R. Lulu, Vanshri Kurhade, Neha Pande, Rohini Gaidhane, Nikita Khangar, Priya Chauhan, Ashwini Khaparde, "Web Camera-based Mouse Controlling Using RGB Colours," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Vol. 4, Issue 2, Feb. 2015

[8] Radhika Bhatt, Nikita Fernandes, Archana Dhage, "Vision-Based Hand Gesture Recognition for Human-Computer Interaction," International Journal of Engineering Science and Innovative Technology (IJESIT), Vol. 2, Issue 3, May 2013.

[9] Heera Lal Bhadrecha, "Finger Gesture Control Computer Mouse With Image Processing," International Journal of Electronics & communication technology IJECT, Vol. 6, Issue 1, Jan. – Mar. 2015.

[10] David J. Rios-Soria, Satu E. Schaeffe, Sara E. Garza-Villarreal, "Handgesture recognition using computer-vision techniques," 21st International Conference on Computer Graphics, Visualization and Computer Vision, 2013.

[11] Granit Luzhnica, Elizabeth Lex, Viktoria Pammer. A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove. In: 3D User Interfaces (3DUI); 2016 IEEE Symposium on ; 2016 Mar 19 ; New York : IEEE; p.81-90, 2016.

[12] C. Wang, Z. Liu and S. C. Chan. Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera. IEEE Transactions on Multimedia; vol. 17, no. 1, pp. 29-39, 2015.

[13] Dulayatrakul J, Prasertsakul P, Kondo T, Nilkhamhang I. Robust implementation of hand gesture recognition for remote human-machine interaction In Information Technology and Electrical Engineering (ICITEE); 2015 7th International Conference on Oct 29, p. 247-252, 2015.

[14] Chen Y, Luo B, Chen YL, Liang G, Wu X. A real-time dynamic hand gesture recognition system using a Kinect sensor. In: Robotics and Biomimetics (ROBIO); 2015 IEEE International Conference on ; 2015 Dec 6; New York : IEEE; p.2026-2030, 2015.

[15] Chen WL, Wu CH, Lin CH. Depth-based hand gesture recognition using hand movements and defects. In: Next-Generation Electronics (ISNE); 2015 International Symposium on; 2015 May 4 ; Taiwan, Taipei ; New York : IEEE; p.1-4, 2015.

[16] Suriya R, Vijaya Chamundeswari V. A survey on hand gesture recognition for simple mouse control, In Information Communication and Embedded Systems (ICICES); 2014 International Conference on; 2014 Feb 27; India, Chennai; New York: IEEE; p.1-5, 2014.

[17] Lee DH, Hong KS. A Hand gesture recognition system based on difference image entropy. In: Advanced Information Management and Service (IMS), 2010 6th International Conference on; 2010 Nov 30; Seoul; New York: IEEE; 2010 ; p.410-413