Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/11454
DOI URL: http://dx.doi.org/10.21474/IJAR01/11454

RESEARCH ARTICLE

REAL TIME HAND GESTURE RECOGNITION BASED IMAGE VIEWER

Sujit Wankhade[1] and A. C. Lomte[2]

1. Student, Computer Engineering, JSPM's BSIOTR, Pune, India.
2. Professor, Computer Engineering, JSPM's BSIOTR, Pune, India.

Manuscript Info

Manuscript History
Received: 31 May 2020
Final Accepted: 30 June 2020
Published: July 2020

Key words:-
Human Machine Interaction, Gesture Based Interactive System, Image Viewer, Web Camera, Multiple Intelligences

Abstract

Human-Machine Interaction based study has been done intensively to improve the ways human communication with the machine. The means of communicating with computers at the moment are limited to keyboards, mice, light pen, trackball, keypads, etc. These devices have grown to be familiar but inherently limit the speed and naturalness with which we interact with the computer. This paper proposes to use image processing and gesture recognition to simulate an environment where performing technical activities will become more like "real-world". The system proposes a real time human hand gesture recognition based system which interacts with the machine and processes the image. The classification of the gesture is based on the direction and gesture of the movement of the human hand. The motion of the hand will be captured and then the keyboard press actions are processed. Depending on this the image will move accordingly left, right, zoom in and zoom out.

Introduction:-

Development in the area of human-computer interaction are very encouraging. Progress in human computer interaction field has also opened a wide opportunity for researchers and scholars to develop new technology. Tools such as Microsoft Kinect, Wii Remote, PlayStation Move and PlayStation Eye have opened a new chapter in this field. Current technology trends are towards interacting with machines using gestures. Research on human machine interaction is very important in improving the ways for human to interact with the machine. These concepts are affecting the fields of human-computer interaction. Users might no longer need to use keyboard and mouse as input devices. These developments will help in integrating technological solution in everyday live.

Researchers need to understand the process, the technique which is time consuming. This is because of gesture are chosen based on the significant relationship with the operation that will done. Human body gestures start to take part in the research after the idea of touch-less based interaction is introduced. The process of interacting is one of the biggest challenge researchers face. Interaction using body gestures are much easier for the user to understand, making the learning process since it does not need an intensive training to learn the gestures.

The method used to interact also should be related with the task carried out by the user. Specific tools are needed for each gesture to be received by the machine from the user. There are several ways users can interact with computer which is by using gestures of voice, touch and hand movement. Studies in the advanced areas have been increased with the mission to improve the way people interact with computers. Interaction using body movement is well suited for applications such as virtual reality application.

Corresponding Author:- Sujit Wankhade
Address:- Student, Computer Engineering, JSPM's BSIOTR, Pune, India.

This project is intended to use image processing to simulate an environment where performing technical activities will become more like "real-world". This project is very closely associated with the emerging Sixth Sense Technology. In further advances to this project, the image processing algorithm can be modified to project the image on the palm of the hand, instead of a screen (which is used here), to make it even more realistic and easy-to-use. Sixth-Sense is a wearable gestural interface device that augments the physical world with digital information and let people use natural hand gestures to interact with that information.

This project takes into account the basic concepts of Mathmizing, GUI, Digital Image Processing and Image Acquisition. This project would cover the concepts of Image Restoration, Image Capturing, Image Enhancement, Image Segmentation and Object Recognition techniques in Digital Image Processing. The Sixth-Sense prototype implements several applications that demonstrate the usefulness, viability and flexibility of the system.

Movements of the finger will be captured by the tools like Kinect or Camera and processed. The Data Computing Device will then determine the gesture movement and perform the keyboard or mouse action as per programmed over the image being displayed in the image viewer. Depending on the motion of the finger the
image will be moved left/right, moved out/moved in as per the user needs.

Related Work :

Human and Machine Interaction based applications have an even greater potential for integrating different technologies by supporting social software applications, presenting e-learning materials and content and offering users games and narrative-based social interactions. But the tutors and learners are required to play their own role for the second step which includes designing activities in-world. Limitations of freedom were imposed in these kind of systems The trend for replicating existing real world buildings and institutions has been almost a first step into using virtual worlds. The visualization work is putting the infrastructure in place for creating engaging experiences. Various hand gesture recognition works have been developed by researchers for different application areas including human computer interaction, virtual game control, robotics and others. System for recognition of hand gestures based on self-organization technique developed by Ishikawa, requires the user to wear a data glove for hand segmentation. But the visualization of existing real world spaces is just one part of the process needed to use these spaces effectively for educational purposes. Through their presence as an avatar in the immersive space, the user can readily feel a sense of control within the environments and more easily engage with the experiences as they unfold. The use of virtual environments has been facilitated greatly through the web-based applications that allow us a range of options including sharing documents and files, holding meetings and events, networking and hosting virtual seminars and lectures, running research experiments, providing forums for sharing research findings and meeting international colleagues. Mathmazing, implemented by researchers from Mumbai Institute, for teaching basic arithmetic to primary school is one of the significant works in the human machine interaction world. In this way, Second Life, in common with other virtual world applications, has opened up the potential for users and learners, teachers and trainers, policy makers and decision makers to collaborate easily in immersive 3D environments. For instance Jennings and Collins (2008) examine 170 accredited educational institutions in Second Life. The success, and wide reporting, of Second Life has helped to highlight the wider use of immersive worlds for supporting a range of human activities and interactions, presenting a wealth of new opportunities for enriching how we learn, how we work and how we play.

Objective and Scope:-

Hand Gesture Recognition Based Image Viewer is based on Sixth Sense Technology which will provide a more natural computer interface, allowing people to move images by their hands and also enlarge them. Sixth sense was first developed by Pranav Mistry in 2009. Sixth sense is not so popular because it is quite difficult to understand as presented by Pranav Mistry. He has introduced many applications based on sixth sense technology such as drawing using gesture, data transfer between two system by physical gesture, etc. he used windows platform and implanted in cpp and c# as programming languages but it is still not implemented in market because of its cost. This work will overcome the limitations faced in the earlier researches conducted by many researchers over reducing the time and cost for good and easy interaction between user and the system. Scope of this work will be more precisely towards reducing the use of hardware devices of data inputs to the computers. This will allow user to interact with the system without more use of hardware peripherals. Our aim to make it simple so that anyone can contribute to improve and evolution of this technology. The main objective of the project is to reduce the cost by implementing sixth sense on open source platform, such as Linux Ubuntu operating system, OpenCV image processing library and python programming language.

The Primary goal of this work is to provide easier access to the computer world and its services. This can lead to new interfaces that will allow the deployment of new commands that are not possible with the current input devices. Hence, users are allowed for richer and user friendlier man-machine interaction. It is also now that Computers have been so tightly integrated with everyday life, that new applications and hardware are constantly introduced. This is probably because the existing devices are inadequate. The means of communicating with computers at the moment are limited to keyboards, mice, light pen, trackball, keypads, etc. Plenty of time will be saved. This work gives communication which is similar to natural way of communication between two humans. Since the introduction of most common input computer devices not a lot have changed. these devices have grown to be familiar but inherently limit the speed and naturalness with which we interact with the computer. Vision based interfaces are feasible and at the present moment the computer is able to "see".

The future of "Virtual Reality Technology" is very promising. In future, this technology will bring out revolution by creating such an atmosphere where people need not be present at places of highly polluted and hazardous industrial areas. In such places, if an image is projected then by simple gestures, as in case of a calculator, one can easily and with high speed and. Moreover, this technology can be employed in banks, offices and public places where too many computers are installed, and fast and efficient calculating speed is required. This data after processing will send a command which can be very easily linked up with the present-day microprocessor technology, thereby ensuring a safe, eco-friendly and healthy work atmosphere. They can simply sit at a place and do their task through gestures on the images received from the site through one set of camera panel, which the camera in the office will track and send data to the server wirelessly.

Methodology:-

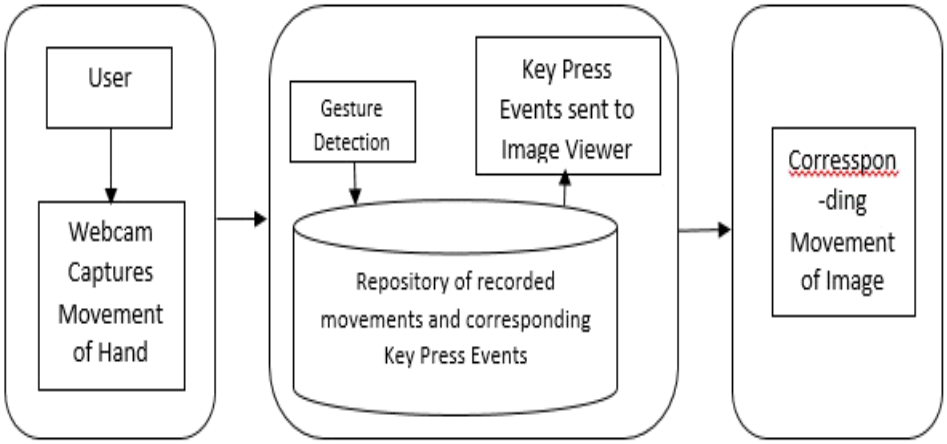


Figure 1:- Workflow.

A webcam captures and recognizes an object in view and tracks the user's hand/fingers gestures using computer-vision based techniques. The camera, in a sense, acts as a digital eye, seeing what the user sees. It also tracks the movements of the thumbs and index fingers of both user's hands. The camera recognizes objects around you instantly, with the micro projector overlaying the information on any surface, including the object itself or your hand. It sends the data to the computing device. Computing software searches the Web and interprets the hand gestures. The movements are then interpreted into gestures that act as interaction instructions for the application interfaces. Depending on the interpretation proper key press actions will be then actioned on the image displayed on the image viewer.

When acquiring an image of a real scene it is discretized in two ways: sampling and quantization. The signal is sample data ten positions (x = 0, .. . , 9), and each sampled value is then quantized to one of seven levels (y = 0, . . . ,6). Sampling and quantization of images is done in the same Way, except that sampling is now done in more than one dimension as shown in Figure 2. Here, we see a continuous signal of two variables (top left), and the corresponding image with signal strength converted to grey values (bottom left). On the right, the signal is shown after sampling on a discrete grid and quantization of the grey values onto five levels. Image acquisition will be done using open cv-python commands.

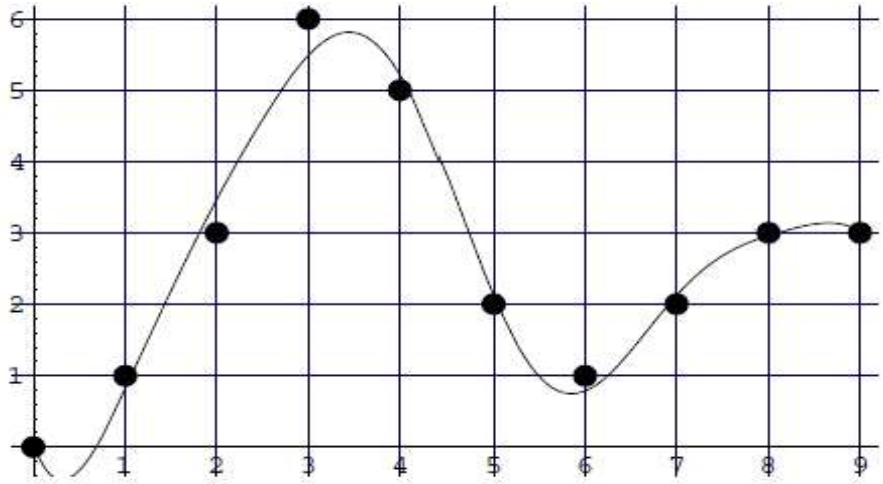
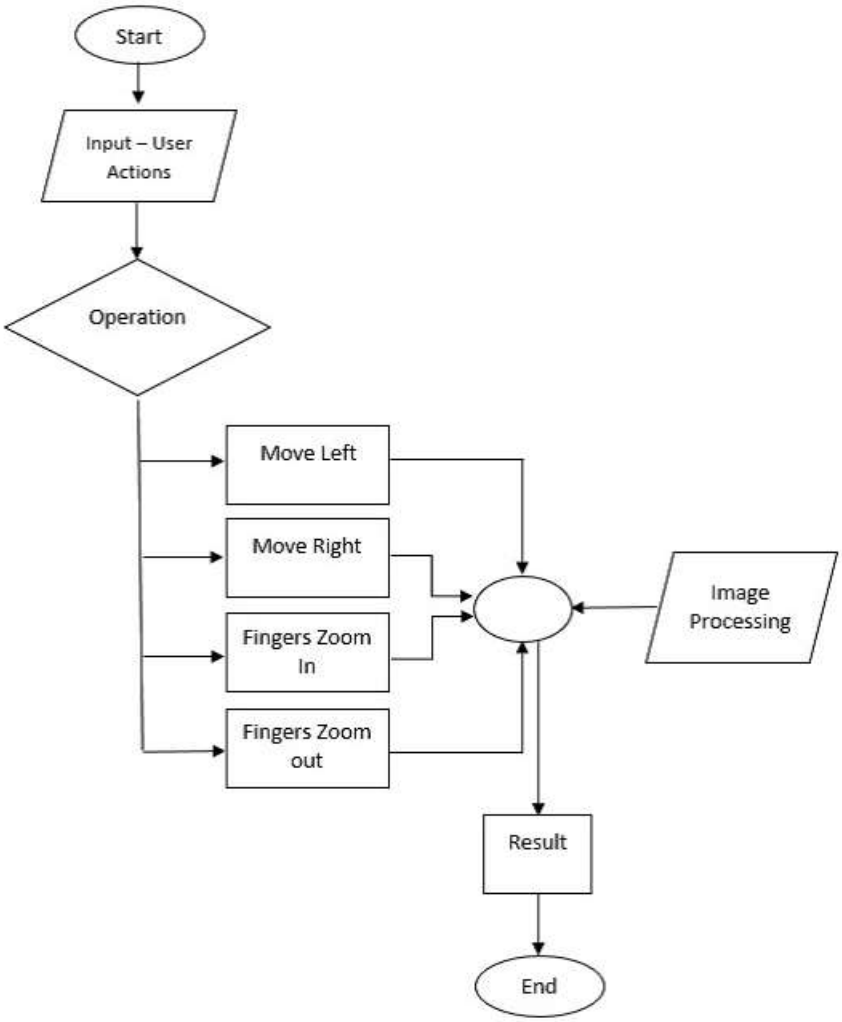


Figure 2:- Image Acquisition.

Actual frame of motion of the hand will be captured and that will be sent to the processing unit. Based on the frame the open-cv - python packages will determine the action to be considered for processing. Packages will then decide the input action to be performed over the picture which is being displayed over the image viewer.

Area of the frame will be detected by counting the pixels diagonally connecting the two detected contours identified and will be divided into several equal parts. Each segment will be dedicated to the key-press actions to be acquired using Image Segmentation.

1. Threshold
2. To make segmentation more robust, the threshold should be automatically selected by the system.
3. Knowledge about the objects, the application, the environment should be used to choose the threshold automatically.
4. Intensity characteristics of the objects Sizes of the objects.
5. Fractions of an image occupied by the objects.
6. Number of different types of objects appearing in an image.
7. If there is no clear valley in the histogram of an image, it means that there are several background pixels that have similar gray level value with object pixels and vice versa.
8. Hysteresis thresholding (i.e., two thresholds, one at each side of the valley) can be used in this case.
9. Pixels above the high threshold are classified as object and below the low threshold as background.
10. Pixels between the low and high thresholds are classified as object only if they are adjacent to other object pixels



Flowchart

Algorithm:

C.1 HSV Segmentation:

The H, S and V separation is done using following equations.

$V = \max \{R,G,B\}$

$\delta = V\text{-}\min \{R,G,B\}$

$S = \delta / V$

Fingers and Palm Segmentation:

1. Capture the Image
2. Read the input image using rad mask
3. Convert RGB image into lab color space
4. Convert the color values.
5. Compute the threshold value.
6. Convert Intensity image into binary image

For Input Image:

Input:

Find the nearest boundary point of one sampled finger point (X, Y)

Step 1. Capture a pixel (x, y) around the sample point of a finger (X, Y)

Step 2. If the value of the pixel is 0 i.e. $P(x, y) == 0$,

goto Step 3. Otherwise, increase angle with a step 1 and then go to Step 1
Step 3. Check the values of 8 neighbours of the pixel (x, y), if it holds $N(x,y)$ is the set of 8 neighbours of the pixel (x, y).

Insert the point into the array of recognition points and capture the movement of points.

Step 4. Capture the next pixel for the movement

1. Continue to search and capture next pixel point of another sampled point until all the sampled points are scanned.
2. Connect all the points recorded in the array of and perform the action accordingly over the image viewer

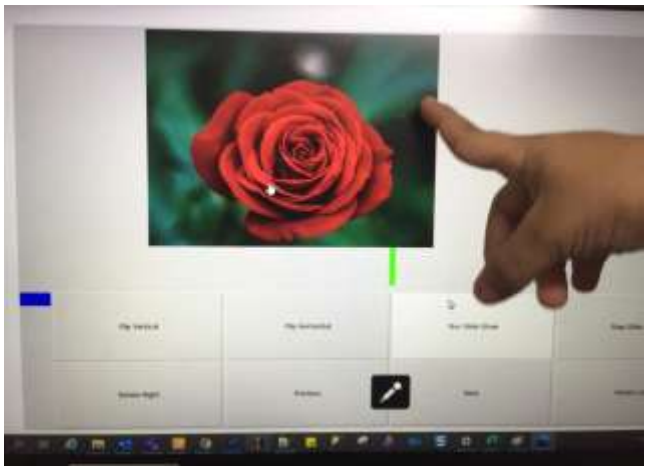
Hand Tracking and Segmentation (HTS) Algorithm:

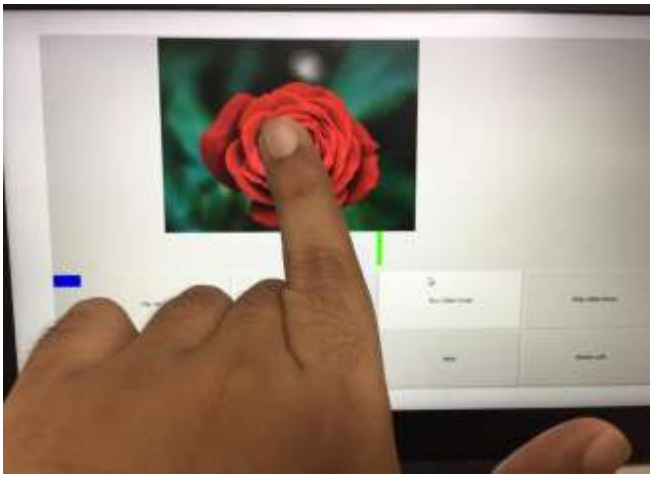
1. Capture the image frames from camera.
2. Process odd frames, track the hand using function by providing skin color.
3. Samples at the run time.
4. HSV histogram is created and the experimented threshold value is passed to the function for tracking required hand portion.
5. Segment the required hand portion from Image.
6. Find the edges by using edge detection.
7. Dilate the Image
8. Image erosion
9. Use the active contours method for simple curve-tracking.
10. Run the Key Press Events

Experimental Results:

Data Sets:

In the trials, two data sets of hand/finger gestures used to evaluate the performance of the system. Whenever a gesture of hand was done, action was performed on the image viewer after capturing the hand gesture movement. All the gesture images belong to females and males. The size of gesture image varies and the image movement also varies.





Time Cost:

Below Table 1 shows the time cost for recognizing the gesture and movement of image accordingly. Unit of Time cost is Seconds. 2nd row shows the gesture performed by hand. 3rd row shows the time cost for the complete process. The experiments are run on the laptop computer of Intel(R) Core(TM) i5-6300U CPU @2.40GHz 2.5GhzGHz and 8 GB RAM and Camera of 0.9MP 16:9 (1280x720) quality.

Sr. No	1	2	3	4	5	6	7	8
Movement	Left	Right	Left	Right	Left	Right	Left	Right
Time Cost	1.1	1.2	1.1	1.24	1.15	1.4	1.1	1.2

Conclusion:-

In this paper we have presented Gesture Based Image Viewer using Image Processing concepts like image restoration, image segmentation and thresholding etc., we could easily make an image viewer which can completely controlled using different human body gestures. Using similar concepts, we can also make various touchless accessories which are indeed need of the hour and can help this world develop in the field of automation.

References:-

1. [1] S. Sahasrabudhe, A. Shah, M. Thakkar, V. Thakkar &S. Iyer, (2012). MathMazing: 3D gesture recognition exergame for arithmetic skills. In Proceedings of 20th International Conference on Computers in Education (ICCE) 2012 (pp. 11-17). Singapore: APSCE
2. [2] Lindley S., Couteur J. and Berthouze N. B. (2008). Stirring up Experience through Movement in Game Play: Effects on Engagement and Social Behaviour. Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems. April 5 10. Florence, Italy. ACM: 511- 514.
3. [3] Nadia Berthouze(2010). Affective posture and body movement as modality for human-computer interaction. Perada Magazine, 10 September, p.1.
4. [4] K. Montgomery, M. Stephanides, S. Schendel, and M. Ross. User interface paradigms for patient-specific surgical planning: lessons learned over a decade of research. Computerized Medical Imaging and Graphics, 29(5):203 222, 2005.
5. [5] Fletcher, J. D., & Tobias, S. (2000) (Eds.). Training and retraining; Handbook of Research on Educational Communications and Technology. A handbook for business, industry, government, and the military. New York: Macmillan
6. [6] Kulik, J.A. (1994). The Impact of Education Technology on Student Achievement. Meta-analytic studies of findings on computer-based instruction. In E.L. Baker, and H.F. O'Neil, Jr. (Eds.). Technology assessment in education and training. Hillsdale, NJ: Lawrence Erlbaum
7. [7] www.blendernation.com/sixth-sense-technology/
8. [8] http://boingboing.net/2009/11/12/sixth-sense-technolo.html
9. [9] http://gizmodo.com/5167790/sixth-sense-technology-may-change-how-we-lookat-the-world-forever
10. [10] http://theviewspaper.net/sixth-sense-technology-will-revolutionize-the-world/
11. [11] http://lucasrichter.wordpress.com/2009/03/13/pattie-maes-sixth-sensetechnology-whats-stopping-this.