Gesture controlled Mouse Pointer for Rehabilitation tools to aid Paralyzed patients

Deepanjali.S, Dr.G. Vadivu
Department of information technology, Faculty of Engineering Technology
SRM Institute of science and technology,
SRM Nagar, Kattankulathur - 603 203, Chennai, Tamil Nadu
deepanjs@srmist.edu.in

ABSTRACT

The aim of the paper deals with controlling the mouse pointer and carrying out the functions of a trackpad with the help of hand gestures. Moving our hand would enable to change tabs, double click an item, drag objects and carry out other simple tasks. Most of the prevailing work uses, an external camera for better image capture and detection which increases the battery power and additional hardware cost. The Deep Neural Network based methodology used in this paper would greatly help partially paralyzed or unwell people due to its increased sensitivity and precision in capturing the gestures even in subdued Lighting.

Key words: Human Computer Vision, image processing, Gesture Controlled, Rehabilitation Tool.

1.Introduction

PC innovation has massively grown over the previous decade and has turned into a vital piece of ordinary life. The essential PC hardware for Human Computer Interaction (HCI) is the mouse. The mouse isn't appropriate for HCI in some genuine circumstances, for example, with Human Robot Interaction (HRI). There have been numerous explores on elective strategies to the PC mouse for HCI. The most characteristic and natural strategy for HCI, that is a reasonable swap for the PC mouse is with the utilization of hand motions. This task is subsequently gone for examining and building up a Computer Control (CC) framework utilizing hand motions. Most workstations today are outfitted with webcams, which have as of late been utilized instability applications using face acknowledgment. So as to tackle the maximum capacity of a webcam, it tends to be utilized for vision-based CC, which would adequately dispose of the requirement for a PC mouse or mouse cushion. The handiness of a webcam can likewise

be extraordinarily stretched out to other HCI application, for example, a communication via gestures database or movement controller.

HCI utilizing hand motions is instinctive and powerful for balanced connection with PCs and it gives a Natural User Interface (NUI). There has been broad research towards novel gadgets and strategies for cursor control utilizing hand signals. Other than HCI, hand motion acknowledgment is additionally utilized in communication via gestures acknowledgment, which makes hand motion acknowledgment much increasingly noteworthy. Human Computer Interface mainly focus on the development of an efficient and easy to use interfaces. The personal computers have a variety of options to interact with different applications efficiently with the use of mouse, track-pad, Joystick etc. Nowadays touchscreen technology is available for devices like mobile phones. But this technology is still costly when used in the personal computers. And the devices currently used to interact are frequently breakable. Our objective was to make an alternative technology to interact with the computer which is not as costly as the touchscreen technology and is easy to operate with in particular to physically challenged person. Contribution of this paper focuses on the following

1. Capability to control the mouse without physically touching the system.
2. A system that ensures vision-based machine with challenges like Low luminance, Distance of the source to the camera.
3. Capability of controlling mouse functions using simple hand gestures for paralyzed patients.
4. To provide real time gesture-based tracking system using Deep Neural Network.

2.RELATED WORK

The recent works that are carried in developing the virtual mouse increases the great impact in conducting the researches in the field of Human Computer Interaction and Computer Vision [18].

The evolution of designing Gesture controlled mouse is organized in the figure below
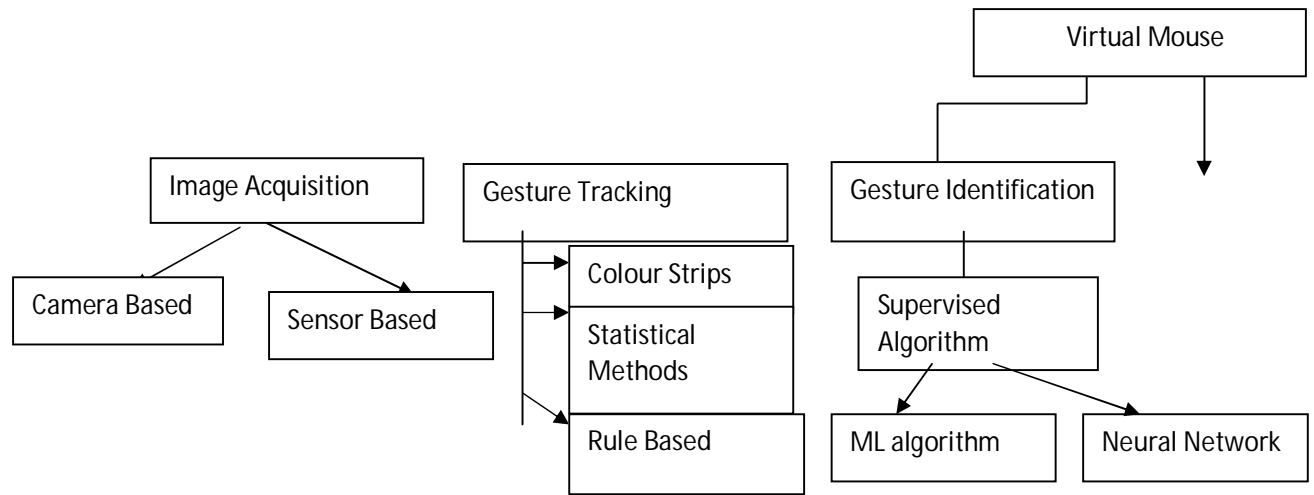
Figure 1: Process Organization of Gesture Controlled virtual Mouse

The capturing of Images can be done via a) Web Camera - which is a vision based or by external wearable's b) Sense Based. The vision Based is a low cost methodology as it uses inbuilt webcam as in many cases [11,12,13,20]. On contrary, The Sensory Based approach makes use of a customized sensing device or wearable's like Wii controller [14], Cyber Glove [15] which is additional cost in development but acquires high accuracy as their not disrupted by external factors like Luminance and distance as in the previous case.

The proceedings of tracking the gesture can be conducted by wearing colour Strips [11,12,13] in the finger this mainly done to show the difference in the hand gestures even at low luminance. The stastical method used in analysing the gestures are Distance metric [9], Convex Hull [16], and centroid measurement of the hand [11,12]. Similarly Rule based methodologies like are also used to track the gestures [1].

The Recognition of the gestures classifies different hand postures which falls under a classic classification task. Hence various linear supervised algorithm [19] like SVM [4] and neural network [4,5,10,17] are used.

3.SYSTEM ARCHITECTURE

The overview of the Architecture is described in the below fig 1. The Main components that were used for the architecture was a 1. Microsoft LifeCam HD-3000(webcam) 2. The computer
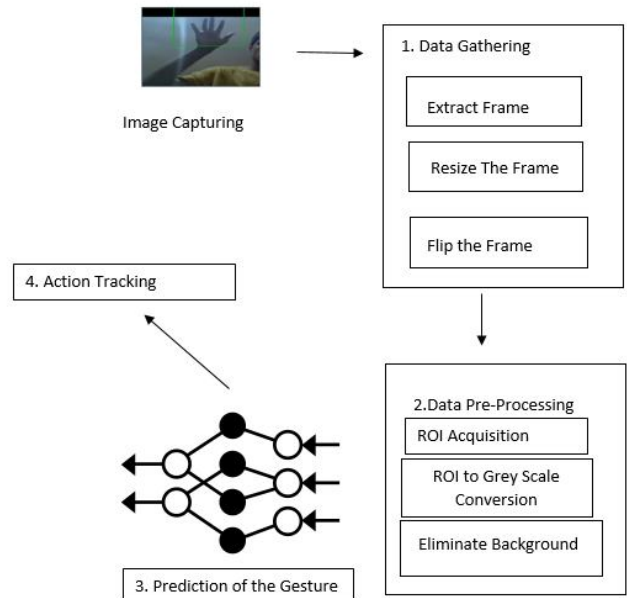
(Intel Core i5-460M, 4 GB RAM). As depicted in the Architecture diagram the webcam is used for sensing the gesture (figure 2). For the purpose of Augmentation, the data acquired is Resized and Flipped. The Data pre-processing phase focuses in eliminating the noise from region of interest and conversion to grey scale. Artificial neural network was used for predicting the class of gestures. Since our methodology included movement based gestures (figure 3) we used libraries to detect the action.

Fig 2: Architecture Diagram

3.1 Data set Gathering

The dataset used in this project is self-generated. For capturing the data, the camera clicks the image in a loop depending on which type of image it is. Every image gets automatically numbered based on the iteration value. The hand images are identified by background elimination technique. First a still background is chosen. And after a few frames, any movement on the screen is considered as the foreground images. The largest foreground image contour is considered as the main hand image so that any small movement does not create any external noise. The hand contour is converted to white colour and the entire background is made black in colour.
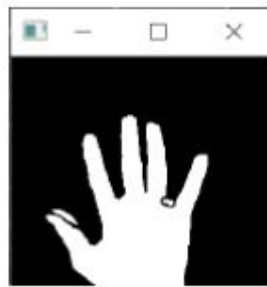
Figure 4: Background Eliminated Image

Once the image is converted to the required format, a keypress begins the recording of the images. One image is recorded in every iteration. Once the required number of images is reached, the loop automatically ends storing every image captured in the hard drive of the local machine. Two additional functions help in background elimination and detecting the required contour of the image. This helps in reducing the noise in the images and provide better image quality.

3.1.1 Extract Individual Frames

In this step, we extract every frame captured by the camera. Since the camera takes the input in the form of a video, we need to convert that video into multiple individual frames. This is necessary because all the computations are done on individual frames and not on the entire video. Once every frame is extracted, the images can be processed.

3.1.2 Resizing and Flipping the Images

When the camera captures the real time feed and converts it into the frames then the frame obtained might be of different size. So, we need to resize the image to fit our need. Now this image is inverted. It is like the image we get when we stand in front of the mirror. So, if we move our coloured strip to right it will in turn move to left and so will our

mouse pointer. This will create a lot of confusion in operating the mouse pointer. Hence, we need to flip each frame. This is done by vertically inverting the frame.

Now this image is inverted. It is like the image we get when we stand in front of the mirror. So, if we move our coloured strip to right it will in turn move to left and so will our mouse pointer. This will create a lot of confusion in operating the mouse pointer. Hence, we need to flip each frame. This is done by vertically inverting the frame

3.2 Pre-Processing Dataset

Once the dataset is generated, all the images must be resized to a fixed size for all the computations to take place. Same sized images are required for all the computations to take place. It eases the computation process if the size of the images is known. Every image is converted to 100 x 89 pixels and the previous image is overwritten with the new resized image

The Region of Interest is a small part of the entire frame which is the actual area where all the processing will be done. The remaining part of the frame is ignored. Only the image inside this region is processed. This is done to remove any unnecessary objects in the background which can cause noise.

The image is converted from an RGB image into a grayscale image so that the foreground and background images can be properly differentiated. This helps in easier computation as there is a clear distinction between the hand and the background. This also helps to remove any noise from the image.

Figure 5: Region of Interest

3.3 Model Training using Neural Network

In Machine Learning, Deep Neural Networks (DNN) is a supervised learning algorithm where the machine recurrently performs the same task on every element of the sequence of where the output of each task is dependent on every previous calculation. This method learns in a sequence called feature hierarchy. Where the features on the

top of the hierarchy are computed by the help of features at the bottom of the hierarchy. DNN's are formulated of multiple layers. The computations in each layer are hidden hence the layers are termed as "hidden layers ". Epoch is a parameter that determines the number of times the algorithm will work on the entire training dataset. In this case, the Epoch is set to 50 meaning the system will repeat the same process 50 times
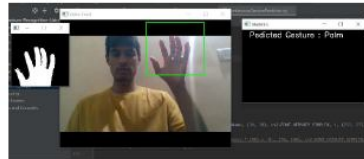

Figure 4.8.1 Palm gesture detected

Figure 4.8.2 Fist gesture detected

Figure 6: Gesture Class Detection

3.4 Action Tracking

3.4.1 Draw the contours and its centroid

The contours are drawn around the hand gesture. Contour is the outer line around a certain area. After this we calculate the centroid. It is the centre. point of the area surrounded by the contour. We control our mouse pointer movement with respect to the calculated centroid.The following are the different actions that are predicted by our proposed system 1. Mouse Movement 2. Left Click 3. Right Click 4. Double Click 5. Scroll Up 6. Scroll Down.

3.4.2 Mouse Movement

This function helps the mouse cursor to move in accordance with the movement of the hand while a fist is made. When the fist gesture is detected, the mouse movement mode is selected. When the mouse is kept still for more than 8 frames, then the mouse gets fixed in that position and any movement made by the hand does not make any changes. A specific gesture has to be made in order to make the mouse move again

4. EXPERIMENT AND RESULT

To train the system, a dataset is required comprising of testing and training data. The dataset used in this project is self-generated. The testing dataset comprises of 6 sets of 1000 images each in which each set denotes a different gesture. The training data comprises of 6 sets of 100 images each in which each set denotes a different gesture. For capturing the data, the camera clicks the image in a loop depending on which type of image it is. Every image gets automatically numbered based on the iteration value.The proposed system was implemented using python and its libraries in different degree of luminance measured(lux) using analogical lux meter and at varying distance.

Illumination Degree(lux)	Gesture Made (left Click)	Gesture Captured Correctly	Recognition Rate
High	15	14	93
Low(< 71 x)	15	12	80
Very Low(< 221 x)	15	11	73

Distance from the source(inches)	Gesture Made (left Click)	Gesture Captured correctly	Recognition Rate
12	15	15	100
24	15	15	100
36	15	13	86
50	15	12	80

Table 1 : Analysis of the system in varied luminance and distance
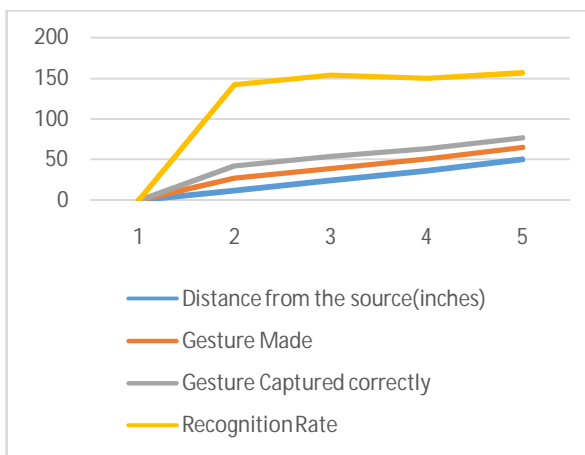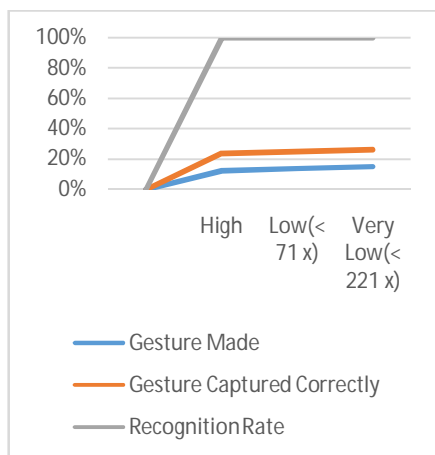
The obtained results analysed in tables and the graph clearly depicts that system has higher
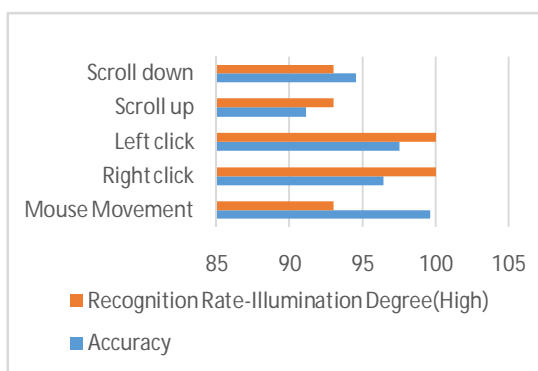
recognition rate in brightened and near environment images. A better background elimination technique to recognize moving backgrounds and eliminate them can increase the accuracy. Sometimes the gesture is not accurately detected, and fluctuation occurs which can cause wrong gestures to be detected. This can be improved by training the system more intensively which would require a very high-performance processor.The system accuracy of the model and recognition rate of optimum environment (illumination Level-High and Distance from source- 12 inches) for various mouse gestures are depicted below

Gesture Type	Accuracy	Recognition Rate-Illumination Degree(High)
Mouse Movement	99.6	93
Right click	96.4	100
Left click	97.5	100
Scroll up	91.1	93
Scroll down	94.5	93

5. CONCLUSION

In this paper we have developed a hand gesture-controlled mouse pointer. This technology has great applications in the fields of computer graphics, gaming, prosthetics, and many more. This technology can be used to help patients who are physically challenged. Most of the earlier applications developed required additional hardware which is often very costly. Our aim was to create this technology in the cheapest all possible way and to create it on a standardized operating system. Various application can be developed using this technology with the minimum requirement of resources. As a future enhancement We would improve the execution of our project particularly the gesture controlling so that it can totally be utilized to supplant our regular mouse. We are also intending to structure an equipment execution for the equivalent in order to improve exactness and increase the usefulness to different spaces, for example, a gaming controller or as a universally useful PC controller.

REFERENCES

1. Liu, K., Kehtarnavaz, N. Real-time robust vision-based hand gesture recognition using stereo images. J Real-Time Image Proc 11, 201–209(2016). https://doi.org/10.1007/s11554-013-0333-6

2. Z. Lu, X. Chen, Q. Li, X. Zhang and P. Zhou, "A Hand Gesture Recognition Framework and Wearable Gesture-Based Interaction Prototype for Mobile Devices," in *IEEE Transactions on Human-Machine* Systems, vol. 44, no. 2, pp. 293-299, April 2014, doi10.1109/THMS.2014.2302794.

3. Luzhnica G, Simon J, Lex E, Pammer V (2016) A sliding window approach to natural hand gesture recognition using a custom data glove. In: 2016 IEEE

symposium on 3D user interfaces (3DUI). IEEE, pp 81–90

4. Marin, G., Dominio, F. &Zanuttigh, P. Hand gesture recognition with jointly calibrated Leap Motion and depth sensor. Multimed Tools Appl 75, 14991–15015 (2016).

5. Neto P, Pereira D, Pires JN, Moreira AP (2013) Real-time and continuous hand gesture spotting: an approach based on artificial neural networks. In: 2013 IEEE international conference on robotics and automation.

6. Rautaray, S.S., Agrawal, A. Vision based hand gesture recognition for human computer interaction: a survey. ArtifIntell Rev 43, 1–54 (2015). https://doi.org/10.1007/s10462-012-9356-9

7. Sharma RP, Verma GK (2015) Human computer interaction using hand gesture. Procedia ComputSci 54:721–727. https://doi.org/10.1016/j.procs.2015.06.085

8. Song Y, Demirdjian D, Davis R (2012) Continuous body and hand gesture recognition for natural human-computer interaction. ACM Trans Interact IntellSyst 2 (1):5:1–5:28. https://doi.org/10.1145/2133366.2133371

9. C. Wang, Z. Liu and S. Chan, "Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera," in IEEE Transactions on Multimedia, vol. 17, no. 1, pp. 29-39, Jan. 2015, doi: 10.1109/TMM.2014.2374357.

10. Xu D (2006) A neural network approach for hand gesture recognition in virtual reality driving training system of spg. In: 18th international conference on pattern recognition (ICPR'06). https://doi.org/10.1109/ICPR.2006.109, vol 3, pp 519–522

11. Grif, H., &Farcas, C.C. (2016). Mouse Cursor Control System Based on Hand Gesture. Procedia Technology, 22, 657-661.

12. Bhadrecha, H.L. Finger Gesture Control Computer Mouse With Image Processing. In Special Issue on AICTE Sponsored International Conference on Data Science & Big Data Analytics for Sustainability (ICDSBD2020)

13. Grif, H., &Farcas, C.C. (2016). Mouse Cursor Control System Based on Hand Gesture. Procedia Technology, 22, 657-661.

14. https://en.wikipedia.org/wiki/Wii_Remote

15. http://www.cyberglovesystems.com/cyber glove-iii

16. Simulation of Mouse using Image Processing Via Convex Hull Method . In International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 3, March 2016.

17. Diliberti N, Peng C, Kaufman C, Dong Y, Hansberger JT (2019) Real-time gesture recognition using 3d sensory data and a light convolutional neural network. In: Proceedings of the 27th ACM international conference on multimedia, MM '19

18. Aigner R, Wigdor D, Benko H, Haller M, Lindbauer D, Ion A, Zhao S, Koh JTKV (2012) Understanding mid-air hand gestures: a study of human preferences in usage of gesture types for hci. Tech. rep. https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci/ https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci/

19. Ahmad al-Qerem1 ArwaAlahmad(2019) Human Body Poses Recognition Using Neural Networks with Data Augmentation, International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, https://doi.org/10.30534/ijatcse/2019/40852019

20. Sukhada Chokkadi , Sannidhan MS , Sudeepa K B , AbhirBhandary,A Study on various state of the art of the Art Face Recognition System using Deep Learning Techniques, International Journal of Advanced Trends in Computer Science and Engineering, Volume 4,