International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

GUI Control with Head Gesture

Reajul Haque Reayz, Prashanta Kumar Dey, Md Kamal Hossain Chowdhury, Mahmudul Hasan

Abstract*— Face detection based mouse movement is one of the important works for the physically challenged people. The movement of head can notify the computer specific task. By understanding, the disabled people computer can pass message to other person and that can help disabled to notify others the urgent need of them. In this project, an application has been built which will help the people who do not have any control of their hands but they can only move their head. In this work, OpenCV based Haar-like feature extraction has been used to detect face. By using face detection algorithm, faces as well as non-faces are classified as faces. So, to reduce the false positive results, face detection algorithm has been enhanced. After face detection, mouse movement has been controlled with the coordinate of face movement.

***Keywords*: Head Gesture, Mouse Control, GUI Control, Computer Operation, Face Movement.**

I. INTRODUCTION

Human Interaction with the digital world makes the globe different. By miniaturization of digital devices, everyone can be easily connected. Devices like keyboard, mouse etc. are used to take input from the individuals. But the disabled person cannot use the devices for their commands. So, face movement or gesture base input system is needed for the person who cannot move their hands. By using the application, they can control the system and it will help them for their needs. Moreover, face detection and tracking are important to understand the user movement. Due to development of wide range of application in security, gaming, entertainment industry, psychological facial expression analysis and human computer interaction, research on face tracking and understanding has been improved. In this project, to help the disabled persons, a virtual mouse is developed to perform like the traditional mouse. By making face movement and gesture control, any user can do functions of an ordinary mouse. The user can do mouse activities like cursor movement, left click, right click, or double click by moving face. The system use Image processing technique to understand and interpret the command from the user.

Manuscript published on 30 September 2019
* Correspondence Author
Reajul Haque Reayz*, Dept. of CSE, Comilla University, Cumilla, Bangladesh. Email: reayz77@yahoo.com
Prashanta Kumar Dey, Dept. of CSE, Comilla University, Cumilla, Bangladesh. Email: proshanto551@gmail.com
Md Kamal Hossain Chowdhury, Asst. Prof., Dept. of CSE, Comilla University,Cumilla, Bangladesh. Email: kamal.cou@gmail.com
Mahmudul Hasan, Asst. Prof., Dept. of CSE, Comilla University, Cumilla, Bangladesh. Email: mhasanraju@gmail.com

II. LITERATURE REVIEW

A variety of application has been developed to detect the human faces and understanding the facial expressions. Human face detection can be classified as an object detection problem. Different methods can be used to detect the object. One of them is Viola Jones method. It proposed a real time object detection framework. Their proposed framework minimizes the computation time as well as achieves detection accuracy. Haar-like features has been used to evaluate quickly the image [1]. Sharif et al proposed [2] Garbor feature method is a useful filter to detect the face. By analyzing specific frequency around the point or region of analysis the object is detected. Category-level object recognition in computer vision can be implemented by constellation model which is a probabilistic, generative model. By setting N parts of images under mutual geometric constraints an object class has been classified. Statistical shape theory based Gaussian derivative filter has been proposed by Burl et al. [3] to detect the features of faces.

III. LIMITATION OF THE EXISTING SYSTEM

There are some limitations on the existing system. They are as follows:

a) For scale normalization, low level preprocessing is necessary because some methods are very sensitive to scale.
b) In some systems, algorithm learning or system learning is very time-consuming, which makes it difficult to evaluate on the face.
c) Many systems are reported a face as non-face images, because those face has unacceptable variation of pose.
d) Many systems take only the fontal part of a face. Face in different angles and different poses are not considered as a standard image input.
e) In some system, preprocessing and detecting a face is more complex than our system.
f) Most of the existing systems have poor accuracy level.

IV. AIMS AND OBJECTIVES

The aims and objectives of the application development are as follows:

a) This system is developed to help disabled people who cannot use their hand. So, by using their body movement, they can control the mouse. It can be said that this system uses face or head as a wireless mouse.

b) Traditional mouse or touch screen based mouse control can be replaced by the developed virtual mouse system which can control the mouse cursor by using web camera. User need to interact with the web camera and by understanding the user interaction, it will control the mouse.
c) The goal is to manipulate the computer with body gestures rather than pointing and clicking a *mouse* or touching a display directly.
d) Try to reduce cost of different hardware.
e) Control and perform mouse operation without touching any device.
f) Human computer interaction also has been tried to enhancement the system.

V. ANALYSIS AND DESIGN

A face detector is a kind of work where an arbitrary size image is given to detect a human face. By analyzing the picture, the system can tell whether the image has human face or not. By using statistical learning methods, face or non-face classifier can be developed. Most of the modern object detection algorithms are based on Viola Jones especially Haar cascades.

Cascaded classifier:
From figure-1, one feature classifier can detect the face with 100% accuracy but 50% of them are false positive. Then it is feed to the five feature based classifier it detects 20% of false positive faces and by using 20 feature based classifier false positive faces reduced to 2%.

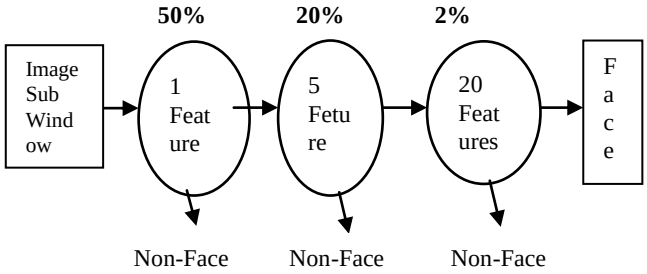

Figure-1: Several classifiers combined to enhance face detection.

Before Haar cascades, adaptive boosting as known as Adaboost was combined several weak classifiers to make strong classifier. By taking several rounds the best weak classifier has been selected and combining the best weak classifiers a strong classifier has been created.

Process of face detection:
Detecting a face in an image is easier than to recognize a specific person face. The general structure of a face needs to determine to understand a face in a certain picture. There is no variation of features in human faces. Like every face has its own nose, eyes,forehead, chin and mouth. All of these combined to detect a face. From the following five figures (Figure 2) it can be understand that they are the parts of a human face.

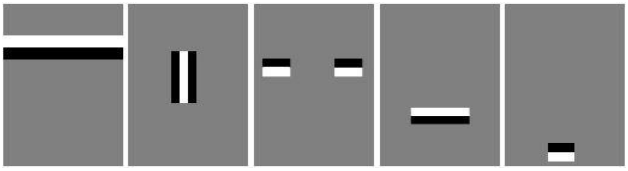

Figure 2: Different figures represent different parts of a face.

By combining the figures, it is a resemblance of a human face (Figure 3).

Combination of figures represents a general human face.

By determining the features in the picture, it can be concluded that the picture has a face or not. If any part of the picture has the partial match, it can say it is a face. It is like the template matching.

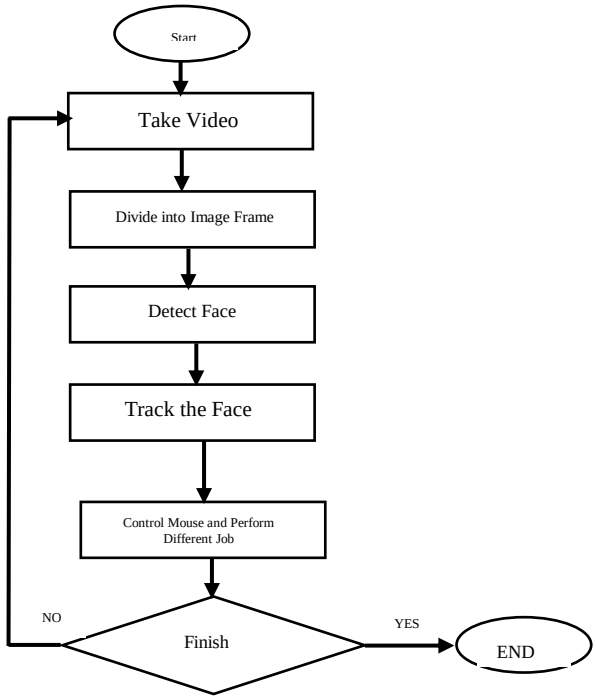

Figure 4: Flowchart of the system.

The proposed procedure:
Our System work on those procedure or step:

Step 1. Take video from web camera.
Step 2. Divide this video into image frame.
Step 3. Detect face on this image using Haar Cascaded feature.
Step 4. Track the face into different image frame.
Step 5. Control mouse position and movement with this face co-ordinate.
Step 6. Perform different operation (like left click, right click, double click).

VI. IMPLEMENTATION AND RESULT DISCUSSION

Installation:

In the system design and implementation, python 2.7.10 and OpenCV 3.0.0 has been used. In the 64-bit windows 10 operating system the project has been implemented.

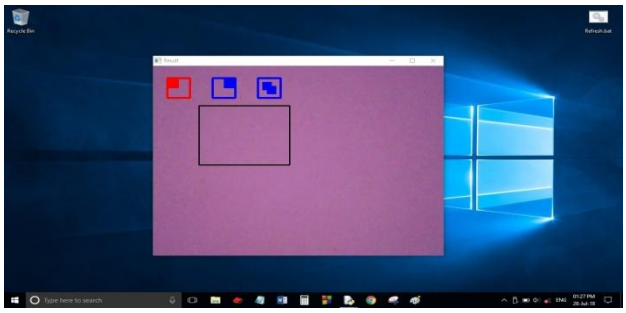

Figure 5: User interface of our system.

In this figure there is a window in middle point. It is denoted our system and take a still picture from a running video. Video has been captured by webcam. The black box or rectangle in the middle is use to control the mouse position. When any figure is find in this box, which will be worked as a mouse cursor. If there are more than one faces in that box, the application choose the largest face among them. It measures the height and width and calculates the area, by using this method the largest face has been detected.

There are also three green boxes in top of our system. Then perform different mouse operation. Such as left click, right click and double click. Those boxes are denoted current status of mouse operation. Left most box is stand for left click status, right most box is stand for double click status, and middle box is stand for right click status. There are also meaningful icons drawn in that box.

Face detection:

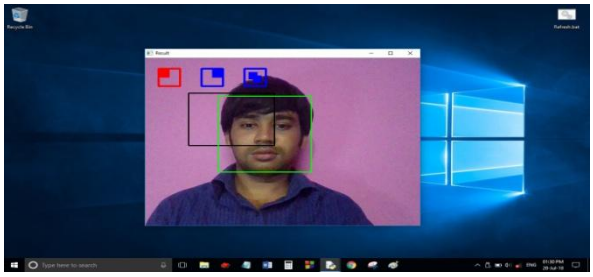

Figure 6: Face detection by our system.

If any face is found in the video frame, then it shown and denoted by a green rectangle. The height and width of this green rectangle is calculated from height and width of shown face. In different frame of video if face is move from its position then this green rectangle also move respectively. This can be called as tracking the face. If there are more than one faces in current frame, then largest face is consider to control the mouse cursor.

Controlling mouse:

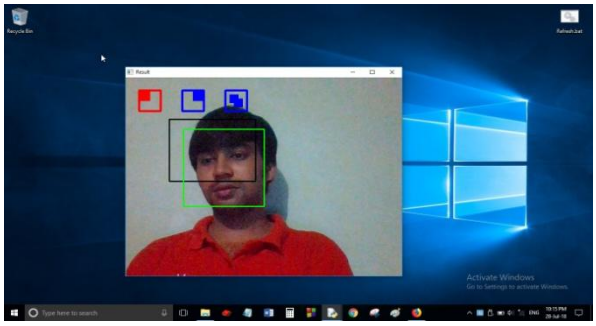

Figure: 7(a)

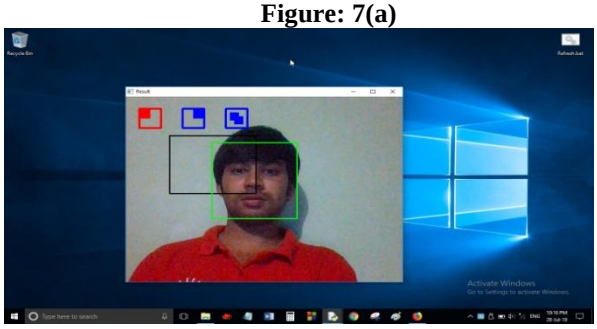

Figure: 7(b)

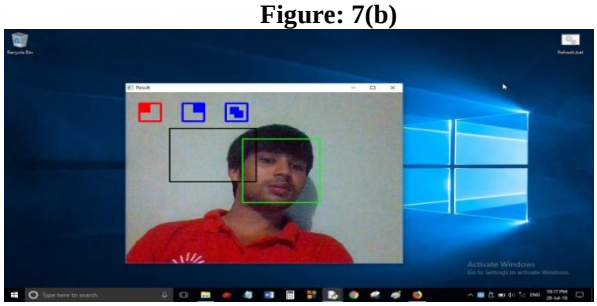

Figure 7(c)

Figure 7: All the images shows the ways how mouse cursor has been controlled in our system.

After face detection phase, the mouse pointer can be controlled. Mouse position is change with the position of face. Notice that there is a black box in our system. If a face detected in upper left corner of this black box like figure 7(a), then cursor will be move to upper left corner denoted (0,0) of display screen. And if face is located at bottom right corner of the black box such as figure 7(c), then cursor also move to bottom right corner of the display screen. The cursor move left to right or top to bottom with a ratio of two different size of display screen and black box size.

Perform mouse operation:

Left click: If anyone wants to perform left click, he/she needs to move face to the leftmost box and change the status. If color of the box is blue, left click in not active. If color is red, left click option is on. It means that if the user puts the face upper left corner of green box and stay more than 3 seconds, it has performed a single left click. To understand whether faces are moving or not, the application considers more than 30 pixels movement.

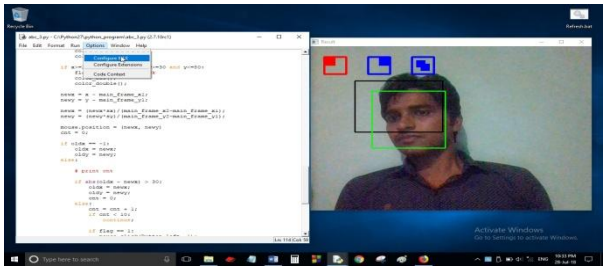

Figure 8: Perform left click.

Right click:

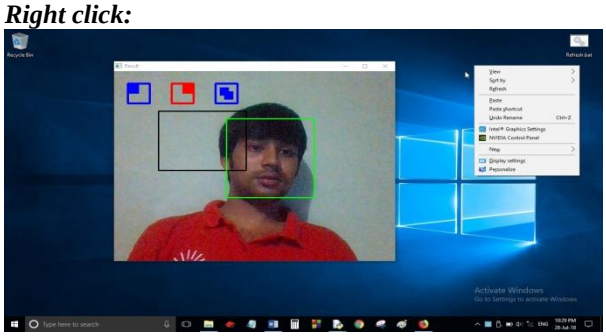

Figure 9: Perform right click.

Again, to perform right click, face needs to move in the middle box and change the status of the mouse. If color of the middle box is red, right click option is on. From the figure 9, it is denoted that a single right click has been performed by upper left corner of green box. For calculating the mouse movement, the application uses the absolute difference between original co-ordinate and current co-ordinate.

Double click: To perform double click, users need to move face to the rightmost box and change the status. If color of the box is blue, double click in not active. If it is red, double click option is on. In the figure 10, a double left click has been performed by upper left corner of green box.

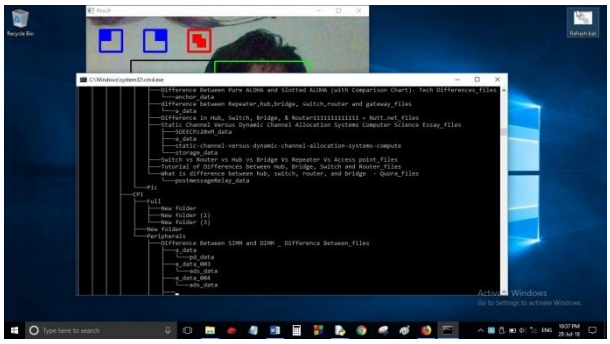

Figure 10: Perform double click.

VII. Result and Discussion

In the developed system, left click operation performs most correctly. In first stage, 15 gestures have been tested and the system can recognize 14 gestures. It means that 94% accuracy level. In the second stage, 15 gestures has been tested and system recognized 12 gestures with 80% accuracy level. Moreover, right click operation accuracy has been tested. Among the 15 gestures of right click has been tested and the system recognized 13 right click gesture. This means that accuracy level is 87%. After that 15 gestures of right click and system recognized 12 gesture with 80% accuracy level.

Table 1: Accuracy rate.

Left Click		
Gesture made	Gesture Recognized	Success Rate
15	14	94%
15	12	80%
Right Click		
Gesture made	Gesture Recognized	Success Rate
15	13	87%
15	12	80%
Double Click		
Gesture made	Gesture Recognized	Success Rate
15	12	80%
15	11	74%

At last, 15 new gestures for double clicking have been tested and the system recognized 12 gestures correctly with 80% accuracy level. Again, 11 new double click gestures were tested and system recognized 11 gestures.

Overall accuracy level of the system is 82.5%. The values are plotted in a bar chart. This chart is shown like this:

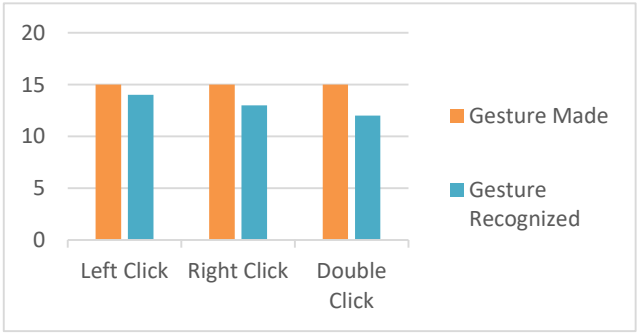

Figure 11: Accuracy level.

VIII. Conclusion

In the system, a software application has been developed to control the mouse using the real time camera which detect the face and does different kinds of operations such as cursor movement, left, right and double click. The system has been developed based on computer vision algorithms. If the computer vision algorithms work well in all environment, the system will work efficiently. However, face detection algorithms faces difficulty in different lighting condition and skin colors of human races. In future, the researchers can detect eye from the faces and control the mouse using the eyes.

Acknowledgment

We specially thank our Computer Science and Engineering Department of Comilla University, Bangladesh for facilitate this research work.

REFERENCES

1. Paul Viola, Michael J. Jones, Fast Multi-view Face Detection, Mitsubishi Electric Research Laboratories, TR2003-096, August 2003.
2. Lawrence Sirovich and Michael Kirby, Low-dimensional procedure for the characterization of human faces. Josa page 519-524, 1987.
3. X. Wang, T. X. Han, and S. Yan, A human detector with partial occlusion handling. In 2009 IEEE 12th International Conference on Computer Vision, pages 32-39, Sept 2009.
4. M. Liu, T. Huang, "Robust Multi- View MultiCamera face detection internal Smart Rooms Using Spatio-Temporal Dynamic Programming," International Conference on Automatic Face and Gesture Recognition, pp.407-412, 2006.
5. Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision 57(2), 2004, USA.
6. Paul Viola, Michael J. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features , Conference on Computer vision and Pattern recognition 2001, USA.
7. H. Rowley, S. Baluja, and T. Kanade, “Rotation invariant neural network-based face detection” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 38- 44, 1998.
8. Rainer Lienhart and Jochen Maydt, An extended set of haar-like features for rapid object detection in Image Processing, on 2002 International Conference, volume 1, IEEE, 2002.
9. A. Erdem, E. Yardimci, Y. Atalay, V. Cetin, A. E. “Computer vision based mouse”, Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASS). IEEE International Conference.
10. Robertson P, Laddaga R, Van Kleek M, International Conference, Virtual mouse vision based interface2004.
11. Jesorsky, O., Kirchberg, K.J., Frischholz, R. W.: Robust face detection using the Hausdorff distance, Lecture Notes in Computer Science. Procs of the Third International Conference on Audio- and Video-Based Person Authentication 2091, 90–95 (2001).
12. Kruppa, H., Castrillón Santana, M., Schiele, B.: Fast and robust face finding via local context. In: Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), pp. 157–164 (2003).
13. Beumer, G., Tao, Q, Bazen, A, Veldhuis, R. A landmark paper in face recognition. In: 7th International Conference on Automatic Face and Gesture Recognition, vol. 78, 2006. FGR 2006, IEEE Computer Society Press, Southampton (2006).
14. Pentland A. Looking at people: sensing for ubiquitous and wearable computing. IEEE Trans. Pattern Anal. Mach. Intell., 22(1), 107–119 (2000).
15. Turk M. “Computer vision in the interface” Commun. Assoc. Comput. Mach. 47(1), 61–67 (2004).
16. Li S.Z. Zhu L. Zhang Z. Blake A. Zhang H. Shum H. “Statistical learning of multi-view face detection: In European Conference Computer Vision, pp. 67–81 (2002).
17. Schneiderman H. Kanade T. “A statistical method for 3d object detection applied to faces and cars” In IEEE Conference on Computer Vision and Pattern Recognition, pp. 1746–1759 (2000).
18. Castrillón M. Déniz O. Antón L. Lorenzo J. “Face and facial feature detection evaluation” In Proceedings of the International Joint Conference on Computer Vision and Computer Graphics Theory and Applications, VISIGRAPP, pp. 167–172 (2008).
19. Lienhart R. Kuranov A. Pisarevsky V. “Empirical analysis of detection cascades of boosted classifiers for rapid object detection” In DAGM’03, 25th Pattern Recognition Symposium, Madgeburg, Germany, pp. 297–304 (2003).

AUTHORS PROFILE

Reajul Haque Reayz Completed his Bachelor of Engineering degree from the department of Computer Science and Engineering of Comilla University, Bangladesh. He was a top level ACM coder and participated many international programming contests.

Prashanta Kumar Dey completed his bachelor degree from the sane department and currently working as a lead programmer in a renowned IT firm. His research interest was on human gesture detection and tracking.

Md. Kamal Hossain Chowdhury is an Assistant Professor of CSE dept. of Comilla University. His research interest is in NLP and Machine learning.

Mahmudul Hasan is currently working as a faculty member (Assistant Professor) in the Department of Computer Science and Engineering at Comilla University, Comilla, Bangladesh. His research interest is in Machine Learning, NLP and Speech Signal Processing