

## Facial Gesture Recognition with Human Computer Interaction for Physically Impaired

Anjusha Pimpalshende<sup>1</sup>, Chalumuru Suresh<sup>2</sup>, Archana Potnurwar<sup>3</sup>, Latika Pinjarkar<sup>4</sup>,  
,Vrushali Bongirwar<sup>5</sup>, Ranjit Dhunde<sup>6</sup>

<sup>1,2</sup> CSE, VNRVJIET, Hyderabad, India

<sup>3</sup>Department of IT, Priyadarshini college of Engineering, Nagpur,

<sup>4</sup>Symbiosis Institute of Technology Nagpur campus, Symbiosis international (Deemed) university, Pune

<sup>5</sup>Shri Ramdeobaba College of Engineering and Management, Nagpur, India

<sup>6</sup>School of Technology Management and Engineering, SVKM'S NMIMS, Navi Mumbai, India

<sup>1</sup>anjusha\_p@vnrvjiet.in, <sup>2</sup>suresh\_ch@vnrvjiet.in, <sup>3</sup>archanapotnurwar@gmail.com <sup>4</sup>latika.pinjarkar@sitnagpur.siu.edu.in,

<sup>5</sup>bongirwarvk@rknec.edu <sup>6</sup>ranjit.dhunde@nmims.edu

### *Article History:*

*Received:* 14-03-2024

*Revised:* 20-05-2024

*Accepted:* 04-06-2024

### **Abstract**

Modern technologies have succeeded in reducing human effort and risks in doing many activities. As physical disability is a barrier for computer interaction based on electronic devices, modern advancements like human-computer interaction (HCI) have proven to be a dynamic field of study in the last few decades which is a major milestone to uphold the lives of the physically impaired in terms of utilizing technology. In this study we projected a method that provides assistance to the physically impaired people for communicating with these techniques. We accomplished this task by considering the facial gestures, assuming that the disability of the physically challenged having no limbs. Distinguished facial gestures, like eye-blinks, mouth open/close, and head position, are captured and fed as input by a camera as a video stream to the operating system. For example, consider a television where a normal human being can perform various operations in it by a remote control with his/her arms but imagine a physically challenged person who cannot interact with the television without human intervention. We used HCI in the form of facial gestures to establish an interaction with the television and physically impaired participants. We incorporated eye-blink as a gesture to switch on/off, mouth open/close to intimate the Operating system to start/stop taking input commands, and the Position of the head is monitored as the input command and is used to alter the volume of the television. An interactive model that recognizes all of these facial gestures provides a great support for the physically impaired people to interact with the systems.

**Keywords:** Human Computer Interaction (HCI), Facial Gestures, Operating System, Human Intervention, Physically Challenged.

## 1. Introduction

Meteorological growth within the evolution of computer technology has led to the idea of human-computer interaction. Although cognitive advancement in technologies is available in the market, not all of them are comprehensible to all the individuals. Physical disabilities can be a barrier for computer interactions based on electronic devices. Visualize a situation where you forgot to take your cell phone to work, you feel your pocket, and suddenly realize that it is not there. It raises an ill-equipped situation and makes you feel that you would never get it back. Empathize with it a million of times worse, within the truth that a person would feel the loss of hands so piercing. Similarly, to watch television, it would be very feasible to simply switch it on with a remotely for a person having their hands. Uncertainly, it would be such a pathetic circumstance for a physically impaired person to switch on the television without an intervention of another human being. Nevertheless, many such instances, like a switching on a fan that requires an action of hands, the case of pulling or pushing a door, typing on a keyboard, driving a car, and most of the times the list is never ending, referring to this complication. However, despite these hindrances, it is entirely the perseverance of the disabled people that they could do their chores in a timely up to an extent without assisting or helping them.

Evolution of technology now a days is a very integral part of routine for every one of us. Modern technologies have succeeded in reducing human effort and risk in doing many other activities. Technology has made us dive deep into it, such that each one of us is glued to various devices and technologies, even in an occupied schedule. Although today's advancement in technology is the culmination of many years of scientific progress and breakthroughs, it is still inaccessible as it is not socially equitable in the society. One such sector of society, being physically impaired, need technology to enhance their lives. Previous WHO estimates that the global disability prevalence is higher and suggested a figure of around 10%. Enabling assistive technologies that includes tools, devices, or software helps improves functions in people who have developed a disability due to accidents, diseases, ageing, or by birth. Many technologies, like assistive cars – a swivel mechanism, mobile app for the deaf, the Q+ app, text to braille converter-a cuboid held with three fingers gets placed on a printed text, brain-controlled mobile application – a device that can be used for reading brain waves using electrodes, gaming for physically disabled, etc. are the some of the latest innovations that have enhanced the routine of the physically challenged.

The motivation for this project is that it is very inconvenient for the physically challenged people to handle mobile phones or any other device to perform basic operations. Our project would provide an easy and feasible access to the physically impaired people to operate the devices effectively through their facial gestures. Human Computer Interaction is one such domain that has provand its ability to elevate the lives of the impaired. It is the study of how people use technological artifacts and interact with the devices in a hands-free environment. User satisfaction is an important facet of human computer interaction as it studies a human's and a machine in communication. HCI seeks to align computer interfaces that constitutes human's mental models, in collaboration with their activities. In the context of humans and the phenomena surrounding them, HCI is a culmination of three concepts: design, implementation, and evaluation of interactive computing systems. HCI can lower the barriers that people encounter disabilities in their routine, such as pulling or pushing a door, switching on a fan, changing channels in a television, and many more. It allows them to avail the benefits of modern

technology, with equal access to sophistication as everyone else. The utility of gestures given a replacement for pivot devices to an HCI. Several types of nonverbal communication (gestures) include body movements and postures, hand gestures, facial expressions, emotions, eye contacts, touch, and voice. Among all these, hand gestures take the chief position while maintaining an interaction between humans and computers. Successful applications of gesture recognition by hands are Human Robot Interaction, sign language recognition, and computer gaming control and many more. Many computer applications can be feasibly operated through hand gestures, among which the most popular ones include the operation of a VLC media player, switching between the websites, changing music in a playlist, or changing slides in MS PowerPoint. Although hand gestures provide adequate features in facilitating efficient interactivity with the computer, it fails to serve the purpose in the case of physically impaired, where the people find it difficult to interact with a computer without hands. This paper primarily focuses on the facial expressions as a chief gesture that are recognized to interact with a device using HCI.

Facial Gesture Recognition (FGR) plays a pivotal role in contributing to the technologies Human Computer Interaction (HCI). Recognition of facial gestures involves capturing of various facial features which can be fed as input that are processed into a sequence of actions and, in turn, gives output to the operating systems. We identified various types of facial gestures, that including eyeblinks, mouth open/close, and head position. Blinking is a facial action that involves a semi-autonomic rapid closing of the eyelid. The Opening and closing of the mouth are identified as a gesture which can be easily detected as a facial action. The Head movements and its positions at different intervals were captured and can be used as a gesture. The above three gestures identified above were integrated into a system which was used to establish a communication between the physically impaired and a computer using HCI.

## 2. Literature Survey

Reza Azad.et.al.[1]have projected an approach that provides an interaction with the virtual environment. A gesture recognition system was implemented which can control the computer media player using hand gestures. The hand and gesture method are applied on theAmerican sign language database with and accuracy is nearly 99.40%, and the planned method can be used in computer games and virtual reality. Papri Ghosh.et.al [2] have projected a holistic approach using HCI to reduce the barrier among thehuman cognitive model and a computer's understanding of a user task. They developed a prototype alongside enhancing the required features and met the needs of desperate users. Xiang Yu [3] has proposed a frameworkby using visual cues that contain head pose detection, face recognition. This study was to find if the automated methods can produce better results when are intercants is truthful or deceptive. The Average detection accuracy has been estimated as 54% and detection of truthfulness was predicted to be 47%. Mahyar Hamedi et.al [4] have proposed elliptic basis function neural network to categorize various facial gestures. EMGs of ten various facial gestures were recorded using 3 pairs of surface electrodes in "bi-polar configuration". With results indicated the best performance to recognize facial gestures was achieved by87.1% accuracy. Thetraining time during classification was estimated as 0.105 sec. Tereza Soukupova.et.al [5] implemented an algorithm to detect eye binks from a standard camera in a video sequence. The Level of eye opening was precisely estimated by detection of the landmarks. This system characterizes eye opening in each

frame. A pattern of ear-based interactions values is detected using a support vector machine classifier in a narrow range. Gloria Zen, Lorenzo Porzi.et.al [6] discussed a framework for personalization classificational models in the paper. The approach was evaluated using various applications like pain recognition and gesture classification etc. The relationship between the distributions and parameters of the corresponding classifiers was achieved through exploitation auxiliary annotated data by the proposed regression framework. Panagiotis Paraskevas Filintisis.et.al [7] has proposed a deep learning method that addressed the in child-robot interaction scenarios. Kuan-Chieh Wang.et.al [8] has proposed a Prototypical Network in the paper “Customizable Facial Gesture Recognition for Improved Assistive Technology”. This solution overcomes the problem of few-shot learning. A novel method was provided that allowed Assistive technology on facial gesture recognition to be customizable using the tracked landmarks as inputs.

Supriya Kapoor.et.al [9] have used “Mahalanobis distance and correlation” to present gesture recognition in real time. “Joy, Anger, Fear, Disgust, Sadness and Surprise” are considered as 6 universal emotional categories. In the training set, there are 50 images consisting size of  $250 * 250$  were used. The results indicated that changing light conditions causes errors, while performance drops with a change in size. Cristina Manresa et.al [10] have implemented an HCI model that is able to track and recognize hand gestures that aims to control videogames. Implemented algorithm can produce best results only when the background is bright. Aashni Hariaa.et.al [11] have discussed a method of the phases of Noise removal and image smoothening, Contour Extraction, Haar Cascade Classifier which detects hand gestures. The proposed method translates the detected gestures into actions, like opening VLC Player and Power point. The Results showed that HCI can be implemented with minimum hardware requirements. Pei Xu.et.al [12] designed an HCI system based on hand gestures. Here components of these methods include hand detection, HCI and gesture recognition. The system recognizes the gestures using Convolutional Neural Networks with a single lens camera and further extended to a Human Robot Interaction case. A simple strategy was developed to avoid false recognition and improve the reliability of system. M Hamed, Sh-Hussain Salleh.al [13] have designed a system for recognizing various facial gestures via neural activities. A versatile interface was designed and implemented in different Human Machine Interaction systems. The Average accuracy was proven to be greater than 90%, could be used as a command controller. Cristina Manresa-Yee.et.al.[14] have projected an interface which identifies facial gestures. Images from the webcam were fed to this interface and were defined as sources of input. X. Zabulis.et.al [15] have implemented a visual-based recognition of hand gestures. The Current state of hand gestures was recorded by typical video cameras. Robots designed to guide who visitors of museum are backed by this approach. Robust human robot interaction is supported by intuitive hand gesture vocabulary and fingertip detection.

Moataz M. Abdelwahab.et.al [16] have presented a novel algorithm explaining the 2D representation of a histogram where a spectrum of angles was applied. The accuracy in this method which is estimated to be 99.17%, can be enhanced by maintaining the spatial relation between the pixels. Javeria Farooq et.al [17] have explored three different techniques for Hand Gesture Recognition (HGR). It was presented as an application. Fingertips were detected by using only a webcam, and the algorithm was developed using MATLAB video processing toolboxes along with computer vision. Jorge García Bueno et.al [18] have implemented a novel algorithm that detects emotions like happiness, sadness, or emotionlessness. A Success rate of 83% was achieved by a humanoid robot that can detect

these emotions. Facial features are determined by Active Appearance Models (AAMs), and emotions are classified using neural evolution and a differential evolution algorithm. MPantic.et.al [19] have proposed an automatic analysis of facial gestures using speech articulation. 22 facial muscle actions were recognized using a robust and flexible method from face image sequence. This paper aims to implement a feature of recognizing the facial gestures depending on whether a person is speaking or not. Three experiments were conducted on facial gesture analysis, which ranging formed AU recognition to silence detection. Md Abdur Rahimetal. [20] have projected an optimal segmentation method to overcome the problems in hand segmentation and recognition because of background illumination, hand size, etc. YCbCr, Skin Mask, and HSV were compared for segmenting hand gestures. To match the color of the hand, segmentation and Skin Mask were combined. Hand gestures were classified using SoftMax classification, and feature extraction was achieved by Convolutional neural networks (CNN).

### 3 Proposed System

In our proposed system, the user interacts with the operating system by providing facial gestures as input commands. These input commands are processed as output actions and displayed to the user. The primary task of this system is to capture the environment, which, in turn, elicits the use of a camera. High-definition cameras and algorithms can be used to handle this process just as accurately and. We have incorporated a camera that captured a video as a stream of frames. In the process of research, we have encountered a challenge i.e. whenever a frame is captured which involves the faces of multiple users, the system fails to detect the real user and hence behaves in an unforeseen manner. To overwhelmed context, we proposed a method for authenticating the user. In the process of authentication, a user must primarily register his/her face with the system beforehand. A set of registered users are stored in the database. As the system encounters multiple faces, only the registered faces in the database are authenticated. To perform authentication facial detection was necessary for authentication

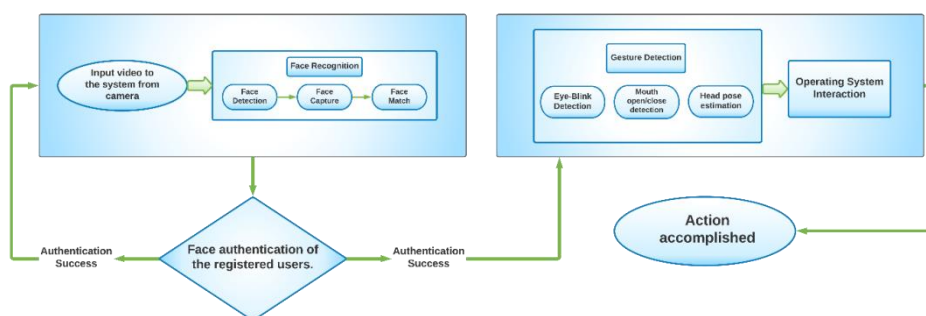


Fig 3.1 Architectural model for the proposed system

Facial recognition plays a crucial role in identifying an individual identity of a person. It involves 3 steps 1. Human faces in images and videos are detected using “Facial Detection”, 2. “Facial Capture” transforms a face (analogue information) to data (digital information) based on facial features, 3. The recognized faces are checked by the “Face Match” if they belong to one of the registered faces. Among all the face detection algorithms invented, Haarcascade algorithm is one of the most powerful facial detections so far. Haar cascade used as an object detection to recognize the faces in an image or a real-time video. A rectangular image is considered, where a simple calculation of the haar value is

done. In the haar feature, the darker areas have pixel values of 0, and the lighter areas have pixel values of 1. Each of these are responsible to find out features in the image or any sudden changes of intensities. To search for a particular facial feature, continuous traversal of the haar feature is done ranging from the top left to the bottom right of the image. Selected features are picked with minimal error rate so that face images can be optimally distinguished from non-face images. Facial authentication deals with one-to-one matching that validates a claimed identity based on a face in an image, either accepting or rejecting a claimed image. Faces that are already recognized or detected by the haar cascades are authenticated by the registered faces in the database of the system. Thus, only the users registered can interact with the system.

### 3.1 Gesture Detection:

Our model detects basic facial features, like eyes, mouth, and head positions, where special facial gestures are captured, like eye blinks, mouth open/close, and head poses.

#### Eye-Blink Detection:

In our model Eye blink detection is divided into two major categories:

Facial landmark detection.

Detecting blinks in video streams.

There are 68 facial coordinates that are estimated by the dlib's library's pre-trained facial landmark detector – "a model built on Histogram of Oriented Gradients + Linear SVM". In a video stream, the frames are looped, followed by resizing and converting them into grey scale. The path of the dlib's library, termed as "--shape-predictor", is used to detect faces in the grey-scale frame.

Facial landmark detection can be applied thereby localizing major regions of face, that including the eyes, eyebrows, nose, and mouth, which implies that special facial gestures can be extracted by indices of face parts. Two sets of facial structures(eyes) were considered to detect the eyeblinks. Starting from the left corner and moving clockwise, each of the eyes are labelled by 6 (x,y) coordinates.

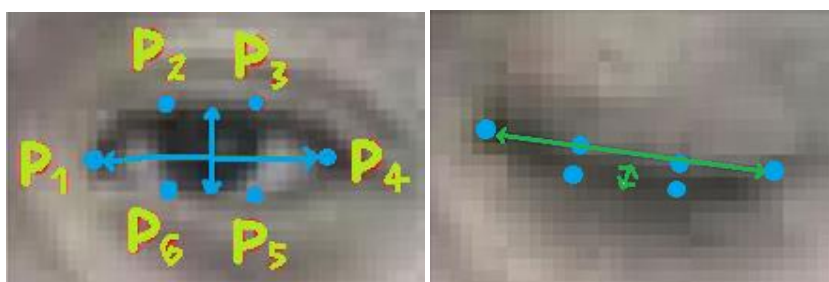


Fig 3.1.1 location of points around the eye from p1 to p6, reduction in the distance of points around the eye when it is closed.

Hence, a relationship is observed within the width and height of these coordinates. This relationship was used to derive an equation: This is termed as Eye Aspect Ratio (EAR).

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

The 2D facial landmark locations are labelled from p1 to p6. Inferring from the above formula, the numerator and denominator are utilized to calculate the distance between “vertical eye landmarks” and “horizontal eye landmarks”, respectively. Since there are “two sets of vertical points and only one set of horizontal points”, the denominator is weighted appropriately. Blinks are detected by this formula in such a way that the EAR is approximately constant, but when a blink occurs, it drastically degrades to zero. To detect a blink, in our model we have considered an EAR threshold of 0.3 where if the ear falls below or raises above this threshold, we considered this as a blink. We examined this threshold for 30 EAR consecutive frames. A blink is considered if the value of EAR is below its threshold limit for all these frames examined. A better blink estimate was obtained by calculating the average of the two Ears’.

Mouth open/close detection:

Another popular facial gesture detection is mouth opening/closing recognition. 12 positional points are obtained in different orders from top and bottom lip feature outputs, as shown in the below fig. 1.2.1.

```
top_lip points: [(181, 359), (192, 339), (211, 332), (225, 336),
(243, 333), (271, 342), (291, 364), (282, 363), (242, 346), (225,
347), (211, 345), (188, 358)]
bottom_lip points: [(291, 364), (270, 389), (243, 401), (223, 403),
(207, 399), (190, 383), (181, 359), (188, 358), (210, 377), (225,
381), (243, 380), (282, 363)]
```

Fig 3.1.2.1 Coordinates of the positional points from top and bottom lip feature outputs.

Matplotlib is used to plot these points. The figure is clearly shown below:

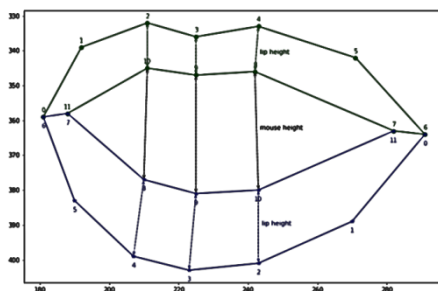


Fig 3.1.2.2 Location of points around the mouth and lips.

From the above figure, the top lip and the bottom lip points were 2-10, 3-9, 4-8. The sum of 12 is always obtained when these indices are added to each point pairs. For example, “i” is considered as one point, and the other is “12-i”. Hence, above is termed as lip height. Similarly, 8-10, 9-9, 10-8 are termed as mouth point pairs. The sum of 18 is always obtained when the indices are added up for each mouth point pairs to calculate the mouth height. For example, “i” is considered as one point, and the other is “18-i”. Hence, above is termed as lip height. Similarly, 8-10, 9-9, 10-8 are termed as mouth point pairs. Therefore, to detect a mouth open/close, the distance is measured as

$$\text{Distance} = \text{math.sqrt}((\text{top lip}[i][0] - \text{bottom lip}[18-i][0])**2 + (\text{top lip}[i][1] - \text{bottom lip}[18-i][1])**2)$$

We used an adjustable ratio of 0.5 to define the extent to which the mouth is open. So, if the height of the mouth is greater than the minimum of the top and the bottom lips multiplied by the ratio (0.5), then a mouth open/close gesture is detected.

### **Head-Pose Estimation:**

As discussed in EAR to distinguish the facial landmarks, the most used dlib's facial landmark detection providing 68 landmarks was applied. For detecting facial landmarks, an annotation box was drawn on the boundaries of the face. For estimating the pose, we considered six points on the face "nose tip, chin, extreme left and right points of the lips, and the left corner of the left eye and right corner of the right eye". The "rotational and translational vectors at the nose tip" are estimated using standard "3D coordinates" of the facial landmarks. After obtaining the required vectors, the 3D points are projected onto a 2D surface which is our image. The angle with the x-axis is found, and with the above-seen annotation box on the face, we can find the intersection of the lines to act as a pointer, along with the angle of movement along with y-axis. Both the results were combined to determine what direction is the head was positioned to recognize the gesture.

#### **3.1.1 Integration with Operating System:**

Finally, when gestures detected, they are fed as input actions to perform tasks in the operating system. We considered a television as our operating system, where tasks like switching on/off television, increasing/decreasing the volume, tuning the channels, etc., are integral operations. Physically challenged people cannot perform these tasks on them without human intervention. The above gestures detected will help physically impaired people to perform these operations. Like,

- **Mouth open/close:** It is used to intimate Operating system to start/stop taking input commands. This is the mandatory operation that must be implemented to ensure that the redundant gestures are not captured.
- **Head position:** The Position of the head is monitored as the input command and is used to alter the volume of the television. The Movement of THE head towards right increases the volume and towards the left decreases it.
- **Eye-Blink:** A single blink of the eye is taken as an input command to switch on/off the television

Hence, the physically impaired people can interact with the television easily without human intervention. Therefore, this action was successfully accomplished.

#### **3.1.2 Experimental Analysis**

Our system was built and implemented using Python. Python libraries like dlib and OpenCV.

Our implemented model captures the facial gestures given by the user, computes required parameters, and performs actions accordingly. Various parameters like computing the EAR, mouth aspect ratio and head pose taken into consideration. Numerous trials were performed with varying threshold parameters a detailed analysis is presented below.



Fig 4.1. Variation of EAR when Eye is open and closed.

A threshold of 0.3 is adjusted for Eye aspect Ratio (EAR) to detect Eye Blink. Three consecutive frames are taken into consideration to confirm that a blink has taken place



Fig 4.2. Variation of MAR when the mouth is opened and closed.

A threshold of 18 as Mouth Aspect Ratio (MAR) was estimated as a perfect threshold by numerous trails for mouth open/close detection.

Frame	EAR	MAR
1	0.463598	0.333333
2	0.345368	4.666667
3	0.339377	8.333333
4	0.288301	15.333333
5	0.115479	19.5
6	0.074085	20.833333
7	0.091784	17.333333
8	0.107143	13.666667
9	0.36873	6.666667
10	0.479964	2.666667

Table 4.1 EAR and MAR values recorded for a single instance of gesture detection.

(One blink and one mouth open/closed)

Table 4.1 is a collection of various EAR and MAR values recorded from the video input for gesture detection. It represents 10 frames with EAR values when a blink is detected. Initially, from frame 1, the values gradually decrease until an eye is completely closed and then when the eye gradually opens, the EAR increases to frame 10. Similarly, in the case of MAR, is at its minimum when the mouth is closed. Then, it gradually increases when mouth starts opening, and at frame 6, when the MAR is 20 units, the mouth is fully opened, and the gesture is detected. It then gradually descended to frame 10 when the mouth was closed.

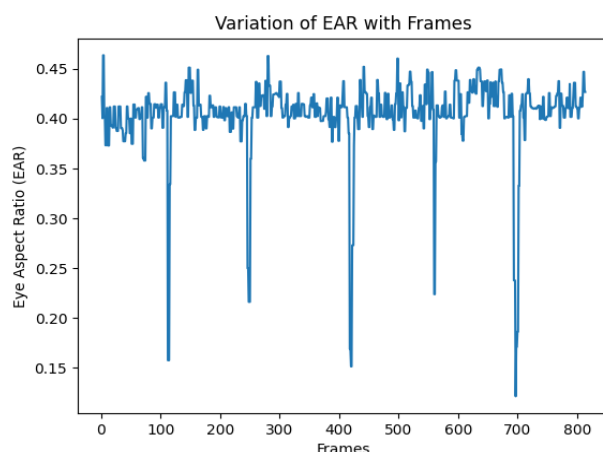


Fig 4.3. Variation of EAR with frames in an experiment of 30 seconds.

We examined the input video for 30 seconds and recorded the EAR for about 832 frames, and plotted the time series graph, as shown in fig 4.3. On analyzing the graph, we observed that EAR value was uniform, standing at 0.40, for most of the time. At this value, no gesture was detected, and a constant activity of the eyes was observed. In the period of 30 seconds, we have voluntarily given five blinks at 3,8,15,20,26 seconds, respectively. These are the intervals where we are experiencing the downfall of the graph by reaching a minimum value of 0.10. We observed that the spikes are sharp, and on examining, we found that blinks happen in fraction of seconds, hence, the graph restores to its normal EAR immediately. After many iterations, we concluded that 0.30 is an optimal threshold value below which a blink is captured.

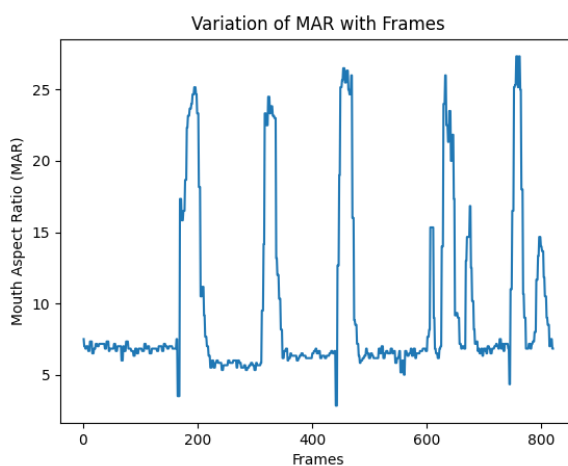


Fig 4.4. Variation of MAR with frames in an experiment of 30 seconds.

Similarly, we have recorded the MAR for 30 seconds for about 820 frames and plotted a time-series graph, as shown in fig 4.4. As shown above, we visualized the graph and on examining a uniform behavior, standing at a value 6.0. At this value, no gestures or activity were detected. In the period of 30 seconds, we voluntarily gave a mouth open gesture five times at 3,9,14,22,27 seconds, respectively. At these intervals, there was a rapid elevation in the graph which can be considered positive for the experiment. Unlike the spikes in the fig 4.3 EAR graph, the spikes here are not too sharp. This is because opening

the mouth and closing it takes comparatively more time than eye blinking. Hence, this spike shows a slightly parabolic behavior. In the fig 4.4, at 600 – 800 intervals, we observed that the graph is not uniform. This phenomenon is due to the normal movement of the mouth. When we talk, our mouth gets a bit opened. We should not consider this as our gesture as these are outliers for this experiment. For this reason, a threshold of 18 was fixed for detecting this mouth-open gesture.

## 5 Conclusion & Future Scope

In our system, we implemented an interactive system by means of HCI that minimizes the barrier between physically incapable and advanced computer technology. Our paper is particularly focused on facial gestures rather than facial expressions, as gestures are the best way to provide accurate operations to the input actions. We have incorporated capturing facial gestures like eye blink detection, mouth open/close detection, and head pose detection, that are fed as input commands to the Operating system which in turn provides output that fosters the routine of physically impediment. The above three gestures identified above were integrated into a system which was used to establish a communication between the physically impaired and a computer using HCI. Unlike traditional methods that use eye localizations or whiteness thresholds of eyes, we have used an efficient method by using the ratio of physical distance between the landmarks of the eyes. In the aspect of mouth recognition, although there are several methods, we chose to apply the simplest where the mouth open height is higher than the height of the lip multiplied by an adjustable ratio up to a threshold, then the mouth is considered as opened. The head poses we detected using 6 points. The required vectors we converted on 2D surfaces to detect the head pose.

In our paper we used television as an operating system, but the interaction is not limited to it. It can be extended to many other electronic devices like switching on/off air conditioners or electric heaters. In the context of facial gestures, the scope can be broadened such a way that many other gestures can be used like vertical nodding of head, winking of single eye, elevating/shrugging the shoulders, and many more. Usage of HCI should be stretched in such a way that evolution of technology should be inclined to physically impaired people. An interactive model that recognizes all these facial gestures provides great support for physically impaired people to interact with the systems. In this context, development of deep understanding of user characteristics and tasks is often emphasized by HCI. Human computer interaction draws from supporting knowledge on both ends of human and machine sides, as it is a study of human and machine in conjunction. Ultimately, HCI proved its potential in adhering to its principles and opportunities in research. It endlessly focuses on providing an interactive approach to every person in society.

## References

- [1] Reza Azad, Babak Azad, Nabil Belhaj Khalifa, Shahram Jamali (2014). "Real-Time Human-Computer Interaction Based on Face and Hand Gesture Recognition", International Journal in Foundations of Computer Science & Technology.
- [2] Papri Ghosh Tejbanta Singh Chingtham Mrinal Kanti Ghose and Ritam Dutta "A Review of Modern HCI Challenges for Betterment of Differently Abled People", International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 9, Number 2 (2013), pp. 79-88.
- [3] Xiang Yu, Shaoting Zhang, Zhennan Yan, Fei Yang, Junzhou Huang, Norah E. Dunbar, Matthew L. Jensen, Judee K. Burgoon, and Dimitris N. Metaxas, "Is Interactional Dissynchrony a Clue to Deception? Insights From Automated Analysis of Nonverbal Visual Cues"

- [4] Mahyar Hamed, Sh-Hussain Salleh, Mehdi Astaraki and Alias Mohd Noor “EMG-based facial gesture recognition through versatile elliptic basis function neural network”, <https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-12-73>, 2013.
- [5] Tereza Soukupova and Jan Cech Center for Machine Perception, Department of Cybernetics “Real-Time Eye Blink Detection using Facial Landmarks”, 21st Computer Vision Winter Workshop Luka Cehovin, RokMandeljc, Vitomir ˇ Struc (eds.) ˇ RimskeToplice, Slovenia, February 3–5, 2016.
- [6] Gloria Zen, Lorenzo Porzi, Enver Sangineto, Elisa Ricci, Member, IEEE, and NicuSebe, Senior Member, IEEE “Learning Personalized Models for Facial Expression Analysis and Gesture Recognition”, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 18, NO. 4, APRIL 2016.
- [7] Panagiotis Paraskevas Filntisis, Niki Efthymiou, Petros Koutras , Gerasimos Potamianos, and Petros Maragos “ Fusing Body Posture With Facial Expressions for Joint Recognition of Affect in Child–Robot Interaction, IEEE ROBOTICS AND AUTOMATION LETTERS, VOL. 4, NO. 4, OCTOBER 2019.
- [8] Kuan-Chieh Wang, Jixuan Wang, Khai Truong, Richard Zemel “CUSTOMIZABLE FACIAL GESTURE RECOGNITION FOR IMPROVED ASSISTIVE TECHNOLOGY”, The ICLR AI for social good workshop 2019.
- [9] Supriya Kapoor, Shruti Khanna and Rahul Bhatia “FACIAL GESTURE RECOGNITION USING CORRELATION AND MAHALANOBIS DISTANCE”, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, Lo. 2, 2010.
- [10] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales “Hand Tracking and Gesture Recognition for Human-Computer Interaction” Electronic Letters on Computer Vision and Image Analysis 5(3):96-104, 2005.
- [11] Aashni Hariaa, ArchanasriSubramaniana, Nivedhitha Asokkumara, Shristi Poddara,Jyothi S Nayaka “Hand Gesture Recognition for Human Computer Interaction” , 7th International Conference on Advances in Computing & Communications, ICACC-2017, 22- 24 August 2017, Cochin, India.
- [12] Pei Xu “A Real-time Hand Gesture Recognition and Human-Computer Interaction System”, arXiv: 1704.07296v1 [cs.CV] 24 Apr 2017.
- [13] M Hamed, Sh-Hussain Salleh, TS Tan, K Ismail, J Ali, C Dee-Uam, C Pavaganun, PP Yupapin “Human facial neural activities and gesture recognition for machine-interfacing applications”, International Journal of Nanomedicine 15 Dec 2011.
- [14] Cristina Manresa-Yee, Javier Varona, and Francisco J. Perales “Towards Hands-Free Interfaces Based on Real-Time Robust Facial Gesture Recognition”, Springer-Verlag Berlin Heidelberg 2006.
- [15] X. Zabulis, H. Baltzakis and A. Argyros “Vision-based Hand Gesture Recognition for Human-Computer Interaction”.
- [16] Moataz M. Abdelwahab, Omnia S. ElSaadany “Face and Gesture Recognition for Human Computer Interaction Employing 2DHoG” 2013 IEEE.
- [17] Javeria Farooq, Muhaddisa Barat Ali “Real Time Hand Gesture Recognition for Computer Interaction”, 2014 International Conference on Robotics and Emerging Allied Technologies in Engineering (iCREATE) Islamabad, Pakistan, April 22-24, 2014.
- [18] Jorge García Bueno, Miguel González-Fierro, Luis Moreno, Carlos Balaguer “Facial Gesture Recognition using Active Appearance Models based on Neural Evolution”, LBR Highlights, March 5–8, 2012, Boston, Massachusetts, USA.
- [19] M. Pantic and L.J.M. Rotbkrantz “Facial gesture recognition in face image sequences: A study on facial gestures typical for speech articulation.”, 2002 IEEE SMC.
- [20] Md Abdur Rahim, Abu Saleh Musa Miah, Abu Sayeed and Jungpil Shin “Hand Gesture Recognition Based on Optimal Segmentation in Human-Computer Interaction”, 3rd IEEE International Conference on Knowledge Innovation and Invention 2020.