International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 09 | Sept -2020 ISSN: 2582-3930

Eye Blink to Voice Communication

Dr. Bindiya M K[1], Rakshitha R[2], Ramyashree U[3], Sirisha R[4], Tejaswini V[5]

[1]*Associate Professor, Department of Computer Science and Engineering & SJB Institute of Technology*
[2,3,4,5]*Student, Department of Computer Science and Engineering & SJB Institute of Technology*
bindiyamk2004@yahoo.com, rakshitha.rajendrag@gmail.com, ramyayadav2016@gmail.com, lishasiri98@gmail.com, tejuyadav1299@gmail.com

Abstract – "Eye Blink to Voice Communication" is a method developed for paralyzed patients used to communicate easily speak with no hurdles. The absolute and complete concern is on the paralyzed or physically disabled patients. Tetraplegia also known as Quadraplegia, refers to paralysis in the upper and lower body, patient cannot move body parts below the neck and they are always directly dependent on someone. The main objective is to style a true concurrent low-cost interactive system that can assist the paralyzed patients through a predefined number of eye-blinks. Haaar Cascade Classifier is appealed for physiognomy and also for eyes to gather the respective information of the patient's eye and full face for smooth functioning. Finally, an Eye-blink to voice or oral meaning it affects both upper and lower body. And also in the majority of cases, there is a chance that the patient can lose his voice because of injury or illness. Due to this condition, the paralyzed communication system which is intended for paralyzed is employed for detecting the eye blinks based on eyelids state(open or close) shall be recorded and that eye blinks are converted into speech or voice form. The system designed exclusively for paralyzed patients which is of low cost so that common people may afford and it's cost-effective so the patient can express their thoughts independently and can lead a smooth life.

***Key Words*:** Haar cascade, eye detection, eye-blink, voice.

1.INTRODUCTION

Tetraplegia is a condition that is caused by injury or illness which ends in a loss of muscle functions and mobility. In other words, they cannot move body parts below the neck. In some cases there will also be a loss of voice. Paraplegic (completely paralyzed) patients require 24-hour support. Hence an effective ocular movement trailing system is introduced. In the recent times thanks to the numerous development of technology there is a real-time considerable request for human-computer and mobile interaction (HCI or HMI). Eye-blinks can be defined as a reflex action of opening and closing the eyelids rapidly. Blink-detection is a prime validated constituent in vigorous domains like human-computer intercommunication, smartphone intercommunication, health community and in many other fields. A concurrent eye blink to voice detection is the evolution of having a feature of blink detection. The suggested structure uses a small part of pupil to trace the actions of the eye. Various techniques are devised for face tracking. Camshaft and Haar Cascade classifier is used for object identification on face in an image or in a video. The chief objective in the present paper is to lodge a structure that can aid the paralyzed in other words developing a real time video oculography system. This can be done by tracing an individual's eye and counts of eye-blink's and those eye-blink counts of paralyzed patients are utilized to communicate independently. Although there are enormous techniques that are being implemented for eye-blink detection, but there is no implementation till now that that is actually being implemented like to put the eye blinks to practical use. The predominant benefaction of this paper is to help the paralyzed and disabled to earn some amount of freedom for communicating independently.

2.OBJECTIVE

The main objective is to develop a real time interactive system that is cost effective for those people who cannot afford. The existing technique for such patients to communicate is too expensive and it is not affordable by everyone.

3.REQUIREMENTS

Eye blink to voice communication gives a very rare form of independence to paralysed patients. The software converts eye blink to voice. So basically, every feature of the software will be accessed and can be controlled by eye blinks only. This software can run on any low-end computers which is affordable by common people as well. The software uses computer vision and Haar cascade classifier which detects the eye blinks and then converts the eye blinks into voice form.

A. Hardware Components

- **Pc or Laptop –** A high speed efficient working laptop.
- **Web camera –** A webcam that is interpolated with 8.0MP still Image and with a clear cut quality of wide-angle lens provides smooth video and permits to enjoy the clarity of webvideo.
- **Power supply –** An efficient power supply for the laptop to keep charged so that it works without any difficulties.
- **Speaker –** To read the voice converted command to read aloud.

B. Software Components

- **Python 3.5.0** – It is a scripting, interpreted language that emphasizes code readability with easy syntax and dynamic semantics.
- **Open CV** – It is a library of python language used to solve computer vision problems.
- **D-lib –** It is a modern C++ toolkit containing machine learning algorithms and tools to solve real world problem. D-lib is very useful in detecting landmarks of facial expression, deep metric learning.

4.METHODOLOGY

Various vision-based eye blinking techniques are diversely used as so many applications such as fatigue monitoring, lie detection, human-computer interfacing[3]. The main purpose of developing the algorithm is to lay out an avaricious way for individuals who can't render it. The main aim is that the system should be affordable for all which includes avaricious constituents for design. The eye-blink detection design consists of majorly 5 steps:-

A. Camera and Frame Capturing

B. Face Landmark

C. Detection of Eye

D. Tracing of Eye

E. Eye blinking

F. Voice conversion

G. Alert

Camera and frame capturing will be done by recording the video of the patient's face and a processed frame procedure creates the frames from the captured video. Face region is used to localize and constitute notable zones of the face such as eyes, eyebrows, nose, mouth, and jawline. Eye detection means recording the action of eye blinks and then the action will be analysed.

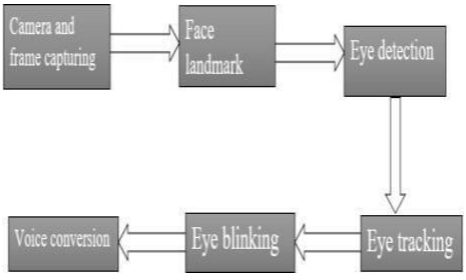

Figure – 1 Venture of the algorithm for eye-blink detection

Eye-tracking is the procedure of measuring where eye look's also known as the point of gaze. All the steps of these algorithms are explained very keenly followed by the scheme. Also, the implementation initiated allows detection of blinks, estimates the duration of eye-blinks, based on those features categorization of eye-blinks can be spontaneous and voluntary.

A. Camera and Frame Capturing

The starting step of the proposed system is initialization. After taking a short video of the patient's face using the webcam and then the processed frame method will be used to create the frames from the captured video. After the frames will be captured the coloured frames will be converted into greyscale only by extracting the luminance component of the eye. The luminosity method is considered to be more sophisticated. This method averages the averages to account for the human eye is further delicate to green when compared to further color so green is most weighted heavily. The luminosity method works best for all.

Figure – 2 Original image and Greyscale image

The starting step of the proposed system is initialization. After taking a short video of the patient's face using the webcam and then the processed frame method will be used to create the frames from the captured video. After the frames will be captured the coloured frames will be converted into greyscale only by extracting the luminance component of the eye. The luminosity method is considered to be more sophisticated. This method averages the averages to account for the human eye is further delicate to green when compared to further color so green is most weighted heavily. The luminosity method works best for all.

B. Face Landmark

Eye-blinks can be recognized by adding on to very particular landmarks of face. Python's d-lib library as already mentioned will be used for facial training sets to understand where certain points exist on facial features.

So after noticing a set of certain facial expressions, facial region detection is therefore a two-step process:

1) Constrain the face image
2) Detect the exterior physiognomy of face

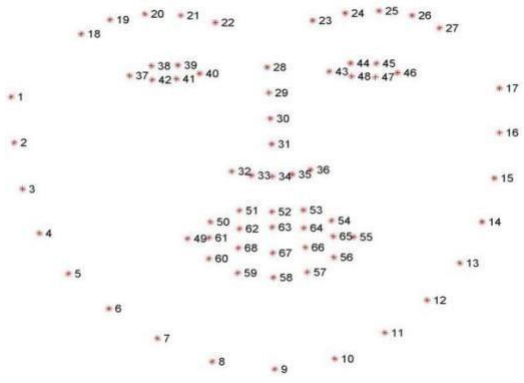

Figure – 3 Dlib regions of face plot

The above figure depicts observations of overall facial

features are expressed in numbers. For eye blinks we have to concentrate from points 37 – 46, in other words, these points describe the eyes.

Facial landmark detection is used to confine predominant areas of face including eyes, nose, ears, and mouth. In terms of blink-detection, the eyes are co-ordinated by 6 (x, y) points, which begins from left corner and then works clockwise around the rest. Considering this image, the key points must be excluded. An association is formed between the width and height of these points Where point1, …, point6 are 2D facial zones. Sokupova and Cech[5] has already derived an equation that represents the Eye Aspect Ratio (EAR). The EAR is an estimate of the eye-opening state. The EAR can be defined by the following equation:

$$EAR = \frac{\|p2-p6\| + \|p3-p5\|}{2\|p1-p4\|} \quad \ldots(1)$$

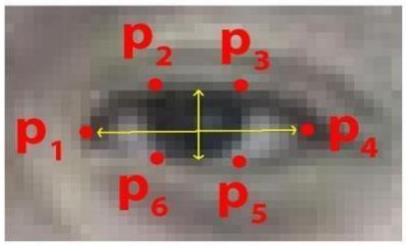

Figure – 4 Figure representing pixel points from p1 to p6 for EAR

Where points p1, p2, p3, p4, p5, p6 represents 2d facial landmark indices. So basically, EAR can be defined as "EAR is a persistent value when the eyelids are open, instantly falls to zero when the eyelids are closed. Figure 5 shows a person's EAR over time.

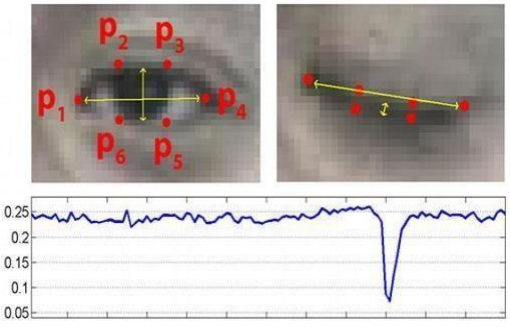

Figure – 5 Facial landmark points for openeye and closed eye and Graph representing Eye values for open and closed eye

Face recognition is a technique to identify or verify the face of a person from a digital image or a video. A human can identify the faces without much effort but when it comes to a computer it is a bit difficult whereas for humans it is an effortless task. There are various complexities such as low resolution, occlusion, illumination, and variations, etc. These factors directly or indirectly affect the computer's accuracy to recognize the face more effectively. First, it is very necessary to understand the main difference amongst face detection and face recognition.

- Face detection –It is generally considered as finding the faces in an image and extract them probably to be used by face detection algorithms.
- Face recognition – The face recognition is used in finding the features that are uniquely described in the image. The facial image is already pulled, sniped, rescaled, and then it will be converted into greyscale. In this case, Haar Cascade classifier is being used (as already mentioned earlier) for face detection and face recognition.

Haar Cascade Classifier: It is one of the machine learning way where a function is trained from both positive and negative images. And then used to detect objects in the face. It mainly has the following steps to configure:

1. **Haar feature selection** - This feature is mainly for being able to detect features in a provided image. The firstly, extract haar features. This considers adjacent rectangular regions at a specific location, in a detection window and then adds up the pixel intensities in all the region and, calculates the difference between the added-up value.

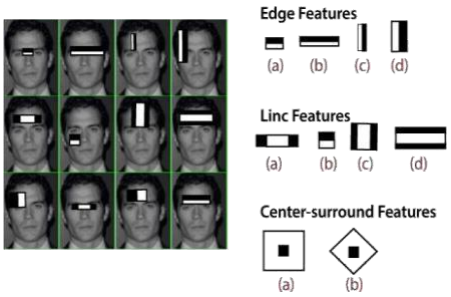

Figure – 6 Rectangular features used for edge feature (haar feature)

2. **Making integral image** - OpenCV uses two approaches for associate object handling search on a unique different scale. One among the approaches is making an associate integral image that's nothing but however considering the standardization maps for a selected facet and so it's recomputed and then the object classifier is applied. The matter with this approach is that once upscaling the $M \times N$ classifier window by 64000$, the options don't map to the pixels grid any longer. To forestall the exact loss that is because of coordinates rounding error, the weights of all options within the classifier cascade have to be compelled to be updated on each scale per the dynamic space features' rectangles as shown in Figure 6.

3. **AdaBoost Coaching** - Eye detection combines and separates scanning results supported an MCT-based Ada-boost algorithm[7]. The formula locates the eyes by putting the attention detector to regions of the eye. AdaBoost that each selects the most effective options and trains them. This formula provides a "stronger" classifier as a combination of weighted

straightforward "weaker" classifiers. Throughout the locating part, a window size of the target is affected on the input image, and for every section of the given image and the Haar options square measures are calculated.

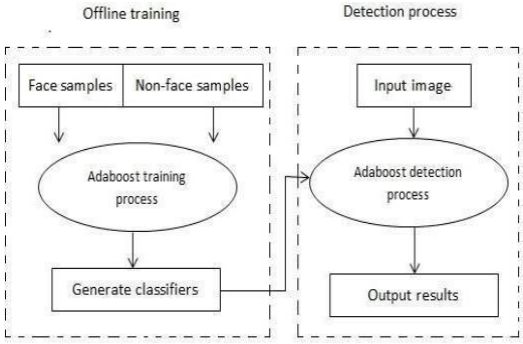

Figure – 7 Flowchart based on Adaboost algorithm

4. Cascading classifier – Cascading classifier is about thoughtful thoughtful about as a robust classifier. This kind of classifier shall consist of a group of stages, where every stage is an associate assemble of weak learners. These learners square measure straightforward classifier known as call stump. Victimization of all these were able to emphasize on the way to face detection happens through the subsequent steps:

   Step -1 The primary step is to load the image in an in an exceedingly greyscale format that's explained very detailed in camera and frame capturing.

   Step-2 When changing the image into grayscale, we are able to do the image manipulation wherever the image will be resized, cropped, blurred, and sharpen if needed.

   Step -3 The Haar-Like feature algorithmic program is employed to seek out the situation of the human faces in frame or image. All the Human faces have some common universal properties of faces just like the eye region is darker than it's neighbour's pixels and nose region is additionally bright than the attention region.

   Step -4 During this step, we have a tendency ro extract the options from the image, with the assistance of edge detection, line detection, and centre detection. Then offer the coordinate of x, y, w, h, that makes a parallelogram hold in the image to indicate the position of the face. It in the desired area where it detects the face as shown in Figure 6.

Face Identification - Face identification is quite easy task for humans when compared to computers. Successful face recognition tends to effective identification of the face features (eyes, nose, mouth, head, face, hairline). The basic idea the recognition of the face is based on the features of a particular face. It is the feasible and most preferred approach for face recognition. The face recognition model can basically operate in two ways: One is Authentication which also indicated for verification of a facial image and the other is Identification or facial recognition. Given below is the detailed steps for determining particular face feature within the face region. (There is a wide range of face landmark detectors however most strategies usually try and localize and then label the subsequent face regions)

a) Face extraction

b) Facial positioning

c) Head create estimate

d) Face switch

e) Blink identification

C. Eye Detection

To eye detection to be done initial, a necessary issue is to own sure haar cascade files to objectify the attention of an individual. Then the options of an eye fixed are actually noted later. So, when it involves eye the very first thing to contmplate is the eyelids. The right, left, higher and lower eyelids are taken initial.

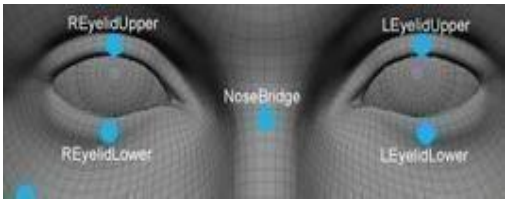

Figure – 8 Figure representing upper and lower eyelids of both eyes.

After eyelids are considered then the next important After eyelids are considered then the next important step is to the extraction of the pupil from the eye. The computer should be ready to represent the attention pupil thus necessary options and also the classifiers are engineered and then applied for sleek functioning.

D. Eye Tracking

Eye Tracking is a method for measuring the motion or movement of the eye which shall indicate the location where the eye is looking at. The main focus here is the movement of eyes can be seen or tracked. Also, the eyeball movement will not lead to any errors. The movement of the eye can be towards any side and this will be extracted, as shown in Figure 9.

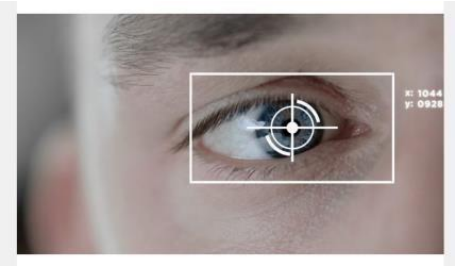

Figure – 9 Image representing eye extraction

Figure – 10 Image representing eye extraction though there is eyeball movement towards any direction.

E. Eye blinking

Eye blinking is majorly outlined because of the gap and shutting of eyelids. The movement of the eye is being observed as well as detected using a highly unobtrusive technique. This mainly is to identify the eye blink which helps in obtaining the voice command, that is to indicate the requirement of the person who's eye blink is being detected.

F. Voice Conversion

Laptop was positioned 60 cm from the patient's face. Blink to Voice software was started which is implemented highly using python programming language. Based on the trained modules appropriate voice commands will be played through the speaker when blinks are detected.

G. Alert through text message

One of the attractive feature of this system is alert will be sent as an text message. This module is achieved through implementation of Twilio application. Twilio allows users to programmatically make and receive phone calls, send and receive text messages, and perform other communication functions using its web service APIs. Each voice command will be interpreted as an text message to the concerned person so that the patient's need can be understood and responded immediately.

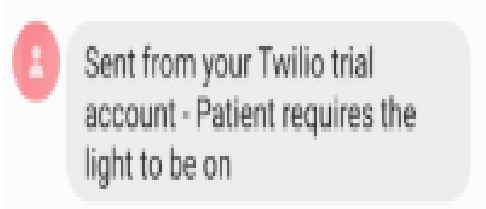

Figure 11 – Screenshot representing text message

5.IMPLEMENTATION

A counter is being assigned for a 'maximum' second and within the threshold value is being checked and counted, and when the maximum threshold value is reached the blink value is being printed and the message assigned to the particular blink value is being converted into the voice and been said. And finally, the counter is reset automatically.

Figure - 12 Snapshots representing Blink to voice software being started for assigned number of counters and is being reset automatically

6.RESULTS

The patient's video is being recorded or captured through webcam where certain aspects such as frames of both eyes and face are detected. These frames are being converted into gray scale. From these frames eye-blinks are detected. The Blink to Voice software will be started where the facial landmark predictor is loaded first and then the video stream thread will be started. As soon as eye-blinks are detected the assigned value for each blink will be sent as an semiotic voice message will be interpreted through speaker. Along with the voice message and alert message will be sent to the concerned person or caretaker.

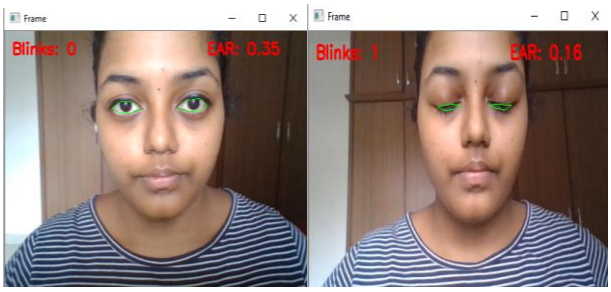

Figure - 13 Image representing eye-blink detection , eye tracing, number of blinks and EAR for each blink

7.CONCLUSION

The proposed system aims to provide an alternative way for paralyzed patients to communicate with the outside world. This proposed system neither harms the patient internally nor externally, as the components of the model will not be in any direct contact with the patient. Hence, it is safer. A useful tool for communication with the machine is an eye blink- controlled USE interface.

Proposed system has the following features:

- Cost-Effective
- Fast
- Accurate

ACKNOWLEDGEMENT

We are very thankful for all the respondents who gave their valuable time and shared their view. This would not have been possible without the guidance, assistance and suggestions of many individuals. We would like to express our deep sense of gratitude and indebtedness to each and everyone who has helped us make this project a success. Last but not the least we would like to thank the Department of Computer Science and Engineering, SJBIT and all our friends who have stood by us and encouraged us in our completion of work.

REFERENCES

1. Sudhir Rao Rupanagudi, Vikas N S, Vivek C Bharadwaj, Manju; Dhruva N; Sowmya K S, "Novel methodology for blink recognition using video oculography for communicating", IEEE Conference Publications, 2014, ICAEE.
2. Masaru Kiyama, Hitoshi Iyatomi and Koichi Ogawa, "Development of robust video-oculography system for non-invasive automatic nerve quantification", IEEE Conference Publications, 2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences.
3. S. M. H. Jansen, H. Kingma, R. L. M. Peeters and R.L. Westra, "A torsional eye movement calculation algorithm for low contrast images in video- oculography", IEEE Conference Publications, 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology.
4. Akshat Gambhir, K. S. Boppaiah, M Shruthi Subbaiah, Pooja M and Kiran P, RNSIT, "Video- Oculographic System using Real-Time Video Processing", International Journal of Computer Applications (0975-8887) Volume 119-No.22, June 2015.
5. Karthik K.P and Basavaraju S, Department of Electronics and Communication Engineering, Sapthagiri College of Engineering, Bengaluru,
6. "Design and Implementation at Unique Video- Oculographic Algorithm using Real Time Video Processing on FGPA", IJSRD-International Journal for Scientific Research and Development | Vol. 3, Issue 04, 2015 | ISSN (online): 2321-0613.
7. Sowmya K S, Roopa B S and Rishika Prakash, Don Bosco Institute of Technology, Electronics and Communications, Bengaluru, "A Novel Eye Movement Recognition Algorithm using Video-Oculography", Proc. Of Int. Conf. on Recent Trends in Signal Processing, Image Processing and VLSI, ICrtSIV.
8. Jacek Naruniec, Michael Wiezczorek, Marek Kowalsak "Webcam-based system for Video- Oculography", 2017.
9. Rachitha S, Prashanth Kumar K. N, "Eyelid movement communication media for paralyzed patient using Fn-M16p module", 2016.
10. S. Saravanakumar, and N. Selvaraju, "Eye Tracking and Blink Detection for Human Computer Interface", International Journal of Computer Applications, Vol 2-No. 2, May 2010.
11. Hewitt. R, "Seeing with OpenCV A Computer- Vision Library," SERVO, January 2007.
12. Jones, M and Viola, P. 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. Computer Vision and Pattern Recognition. IEEE Computer Society Conference.
13. Himani, S. P., Darshak,G. T. and Udesang, K. Jal. 2014. A Survey on Object Detection and Tracking Methods. International Journal of Innovative. Research in Computer and Communication Engineering.