Eye Tracking-Based Communication Assistor for ALS

Agfa Chris[1], Gayathri Suresh[2], Joshua S George[3], Ms. Sajni Nirmal[4]
[1,2]Associate Software Engineer, Marian Engineering College
[3]Data Analytics Aspirant, Marian Engineering College
[4]Assistant Professor, Marian Engineering College

Abstract **- Amyotrophic lateral sclerosis (ALS) is a neurological disease that progressively affects the nerves that control voluntary movement in humans. It is the most common motor neuron disease seen in adults, with around 2 in 100,000 developing ALS yearly. This life-threatening disease results in the loss of muscle movements in affected patients. As a result, patients eventually lose basic capabilities like eating, speaking, walking and so on. Even after decades of medical advancements, there still exists no cure for this disease. In recent times, Assistive Technology (AT) has proved to be an effective way to improve the lives of people affected by ALS, and help them carry out their day-to-day tasks with ease. These include mobility-enhancing and speech-generating devices, but most existing technologies are cumbersome for the patient to use, and require extensive training on their part. In this work, we propose to develop an eye-based communication system that can resolve the speech difficulties faced by patients. The system records and processes the eye movements of the user, which are then translated into corresponding phrases for communication. By the application of image processing algorithms, the system can be used by ALS patients to carry out day-to-day conversations with their near and dear with minimal training and effort.**

Index Terms **– ALS, eye-tracking, communication, assistive technology.**

I.INTRODUCTION

Amyotrophic Lateral Sclerosis (ALS) gained worldwide attention in 2014 when Pat Quinn and Pete Frates, who were patients of ALS, co-created an ice-bucket challenge for spreading awareness about the disease. Their efforts focused the spotlight on ALS patients and their lonely and lifelong battle with the disease. Victims of ALS progressively lose all control over motor neurons, eventually being unable to eat, speak or move. This results in great despair and suffering both for the patient and the caregiver.

In this context, research in this field is concentrated towards developing systems that can reduce suffering and improve the quality of life of those affected. The past few years have seen the development of a number of such models, that seek to help ALS patients regain their lost capabilities, at least to a certain extent.

Though patients of this disease eventually lose their muscular capabilities, the muscles of the eye still retain the capacity for movement. Till the very end of their life, they remain able to control their eyeball movements. Ongoing research in the field of medical engineering is currently focusing on utilising this potential for the development of assistive technologies for ALS. Nevertheless, existing systems are generally expensive, and are quite cumbersome to use for the patient. Thus, we seek to develop an easy-to-use and effective method to provide communication assistance, by harnessing the movements of the eye. An attached camera captures images of the eye, which are then processed to detect and classify the direction of the patient's gaze. These positions are then translated by the system into commonly used phrases.

The remaining paper is structured as follows: Section II de- scribes the already existing systems developed for the purpose of eye-based communication assistance. Section III and IV provide a detailed explanation of the proposed design. Section V describes the design of the system interface. Section VI details the experiments conducted and the inferred results.

II.EXISTING WORK

Over the years, researchers have put forth various designs for the purpose of implementing the process of eye tracking and gaze detection for ALS patients. Considering that accuracy remains one of the most critical features that determine a system's efficiency, many of the existing technologies fail to produce reliable results and lack user-friendliness. The following sections describe some of the eye-tracking technologies that are currently in use.

A.EyeGaze Edge

This device enables users who have lost the ability to speak and move to type out words on a screen using movements of the eye [2]. The user is directed to look at rectangular cells that are displayed on the device screen. To “press” a particular key, the user looks at the key for a specified period of time. The device was conceived as a system that would be of help to writers/authors affected by ALS, as it could help them type out words on a screen.

Limitations: Though particularly useful for the purpose of writing, it would require some significant effort from the patient to type out a single sentence on screen, just for the purpose of conversation. Priced at $8700, it is also quite costly for an average user. Additionally, it uses two screens, which limits its portability.

B.EOG-Based Devices

Electrooculography (EOG) techniques are a popular means of eye-tracking in many devices. It utilises the potential difference between the front and the back of the human eye for recording the position of the eye [3]. For measuring the position of the eyeball, a pair of electrodes are positioned above and below the eye, or to the left and right of the eye. When the eye moves from the center towards left or right, the corresponding electrode is exposed to the positive side of the eye’s retina while the other electrode sees the negative side. Thus there arises a difference in potential between the two electrodes. The resulting signals determine the direction of eyeball movement.

Limitations: Although electrooculography is widely used by ophthalmologists for diagnosis and tracking, image processing techniques are generally considered to be capable of producing more accurate results. This is due to the fact that there are various other parts of the human body that also generate signals, making it difficult to estimate an accurate gaze position based on the EOG signals.

C. Asistsys

Asistsys was designed as a complex system that would facilitate communication between caregivers and patients, and as a way for the caregiver to keep track of the patient. It uses a camera to capture movements of the eyes of the patient, which are then processed to map to certain pre-fixed phrases [4]. These phrases are communicated to the caregiver via an SMS notification system.

Limitations: The system requires the camera to be fitted very close to the eye of the patient, as it cannot detect and extract the eye portions from the entire image of the face. Thus it can be a source of discomfort for the user.

III.PROPOSED APPROACH

The proposed system is designed to be an easy-to-use eye-tracking application that provides basic communication capabilities for patients of ALS. The system methodology operates in five main steps as detailed below:

A. Video Acquisition

The first step towards accurate detection and processing is the acquisition of clear and noise-free images of the object of interest, in this case, the patient. The camera that is attached to the system captures videos of the user as he/she looks at it. These images are passed on to the processing system for further operations.

B. Face Localization

Once we have obtained image data from the capturing device, the next step is to detect the presence of a face in the captured image and to extract the necessary facial features from it [1]. This is generally referred to as face localization. Performing this step eliminates the need for the camera to be head-mounted and thus allows for a more comfortable experience. Image processing can be applied to detect and localize the face of the patient, ignoring other background information. For this, it is preferred to apply the Haar Cascade classifier proposed by Paul Viola and Michael Jones in 2001 [5], which employs machine learning algorithms to train a cascading function using images with and without faces.

In order to eliminate false detections, the algorithm returns only the biggest detected object in the frame. This is based on the observation that the Haar cascade tends to falsely detect other objects in the image, but the biggest of the detected objects is usually the face.

C. Eye Detection

Eye detection is one of the most crucial stages of processing. The accuracy, and ultimately the efficiency of the design depends on how effectively it is able to detect and process the eye movements of the patient. Once the face is localized from the captured image, we use the Haar Cascade eye classifier to detect the eyes from the face frame. The eye points are then separated from the extracted facial points. Since the image of the eye consists of extraneous features like eyelashes, eyebrows and eyelids, we use blob detection to detect the presence of the pupil in the image [6].

In order to optimize eye detection and handle false detections, the function returns only those objects that are detected in the top half of the face frame.

D. Pupil Detection

The possibility of using the Circular Hough Transform as a means of detecting the pupil in the eye frame was initially explored [1], but failed to give accurate results. Thus blob detection has been employed as a reasonably reliable method to detect the pupil of the eye [6]. A Blob is a group of connected pixels in an image that share some common property, distinguishable from its background. Blob detection algorithms use simple image processing like thresholding and merging to detect and extract the blobs in a given image.

E. Gaze Detection

In the fourth stage, we carry out gaze detection, which involves determining the direction of the patient’s gaze. There are a large number of methodologies that

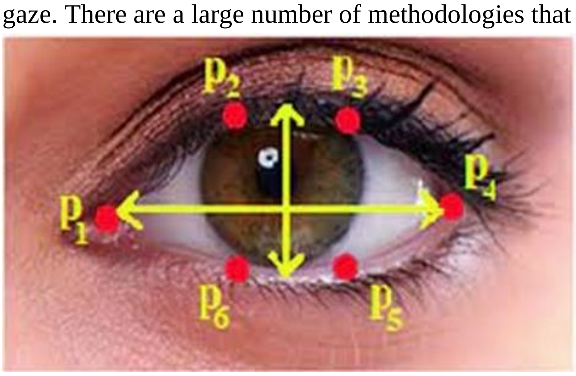

Fig. 1. Six facial landmarks associated with the eye

exist for this purpose [7]. One simple, yet effective technique is to calculate the gaze ratio of the eye. The eye frame is divided into two halves. A left gaze implies that there will be more sclera (white of the eye) visible in the right half of the eye than the left half. Similarly, a rightward-gazing eye will have more sclera visible in the left half than the right half. This information can be used to determine the gaze direction of the eye. The method proceeds as follows:

1)Divide the eye frame into two halves (left and right).

2)Count the number of white pixels in the left half and right half.

3)The ratio of white pixels in the left half to that in the right half gives the gaze ratio.

F. Blink Detection

For the purpose of blink detection, we utilise the Eye Aspect Ratio (EAR), which is an elegant and efficient means of detecting blinks in a video feed [8]. An eye is represented by a set of six (x,y) coordinates P1, P2, P3, P4, P5 and P6 as depicted in Fig. 1. The EAR is calculated as in (1). The numerator calculates the distance between the vertical eye coordinates and the denominator calculates the distance between the horizontal eye coordinates. The EAR remains at a relatively constant value while the eye is open, but suddenly falls to zero when a blink occurs. Thus the algorithm monitors the EAR value, and records a blink whenever the ratio falls below a certain pre-specified threshold.

$$EAR = \frac{||P_2 - P_6|| + ||P_3 - P_4||}{2||P_1 - P_4||} \quad (1)$$

G. Gaze-to-Command Mapping

As the final step, we create a database of predefined common phrases and commands which can be modified or updated by the user. The calculated gaze direction is then mapped to a specific command from the database and highlighted on screen. For example, a leftward movement of the patient’s eyeball can map to ”I’m feeling hungry”, a rightward gaze can mean ”Call the doctor”, and so on. This mapping facilitates communication between the patient and the caregiver.

IV. WORKFLOW

The activity diagram for the proposed system is depicted in Fig. 2, which outlines the five modules of our system. When the application is started, the attached/inbuilt camera begins capturing real time video of the patient. Each frame obtained by the camera is passed on to the next stage of processing,

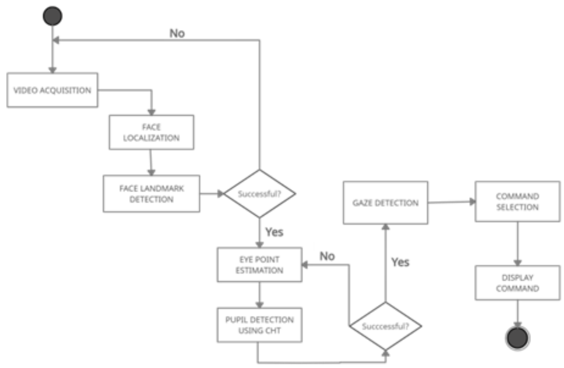

Fig. 2. Proposed workflow of the system

i.e., face detection. Haar cascade-based classifiers are used to detect the face of the patient from the input frames. From the detected face frame, the eyes of the patient are extracted using a cascading function trained on a number of template images. Once we have obtained the eye frames of the patient, the system proceeds to localize the pupil. This is achieved by blob detection, which uses thresholding and merging to detect a group of connected similar pixels in an image. At a suitable threshold, the blob detection algorithm can accurately detect the pupils of the patient. On successful pupil detection, the algorithm calculates the gaze ratio of the eye to determine the direction of the patient's gaze. This is based on the number of white pixels visible in the left and right halves of the eye. The corresponding command is then highlighted on screen based on the calculated gaze ratio.

V.USER INTERFACE

The user interface is designed to be as minimal and user- friendly as possible, thus enabling the patient to use the system with minimum effort on his/her part. In such a scenario, designing the interface is a major challenge as there is no scope for conventional touch/click interactions between the user and the system. The user navigates between the various windows solely by the movement of his/her eyeballs.

The system is activated when the user blinks twice. This displays a screen with various commands that enable the patient to carry out independent tasks like playing music or opening a folder. The patient selects the desired action by looking in the corresponding direction. The user can also choose to ring an alarm, which navigates to a new screen containing phrases the user may want to say. By looking left or right, the user can highlight the desired phrase onscreen.

VI.RESULTS AND DISCUSSIONS

A. Results

The application was tested with users seated at varying distances from the camera, and in various lighting conditions, and was observed to be working as intended. The application was able to detect the direction of the user's eye movement and perform corresponding actions with reasonable accuracy.

It is thus able to fulfil the basic communication needs of ALS- afflicted individuals who are unable to move any part of their body other than their eye muscles. The user can also play and pause music from a playlist by looking in the corresponding direction.

Table I depicts the threshold values obtained in various lighting conditions. The system accurately detected the pupil at a threshold value approximately equal to 42 in good light. It was observed that the iris was accurately detected in 96% of the cases tested in daylight as opposed to an accuracy of 72% in darkness. Similarly, gaze was correctly identified in 20 of the 25 cases tested in daylight. It was noted that the application required the user to be seated directly facing the light source for proper operation.

TABLE I DETECTION ACCURACY IN VARIOUS LIGHTING CONDITIONS (OUT OF 25 TEST CASES)

Lighting	Threshold	Iris Detected	Gaze Detected
Tubelight	42	23	20
Sunlight	51	24	20
Darkness	53	18	15

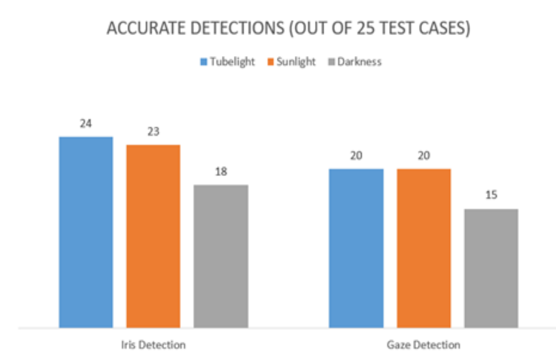

Fig. 3. Accuracy of iris and gaze detection under various lighting conditions

Calculation of the gaze ratio is the crucial step that enables the detection of gaze direction. The gaze ratio values of a sample set of 10 random camera frames is given in Table II and graphically depicted in Fig 4. The algorithm determines the direction of the patient's eyeball movements based on the obtained range of the ratios for each direction.

B. Challenges

During the development of the application we faced various challenges, especially in the detection of gaze. While calculating the gaze ratio, the bottom gaze was comparatively difficult to detect. The algorithm produced many false detections by mistaking the bottom gaze for a closed eye/blink. Due to the shortcomings in existing threshold functions, it is difficult to detect left, right, top and bottom gaze separately. As shown in Table II, the gaze ratio values overlap considerably. Similarly, the application requires an appropriate threshold to be manually set by the user in order to be able to accurately

TABLE II OBTAINED GAZE RATIOS

Gaze Ratios			
Left ratio	*Right ratio*	*Top ratio*	*Bottom ratio*
1.044	0.937	0.714	0.973
1.165	0.842	0.629	0.863
1.069	0.967	0.622	0.933
1.074	0.854	0.645	0.917
1.086	0.974	0.665	0.908
1.120	0.863	0.729	0.922
1.095	0.918	0.632	0.959
1.086	0.872	0.693	0.825
1.003	0.922	0.76	5 0.914
1.055	0.894	0.8	0.937

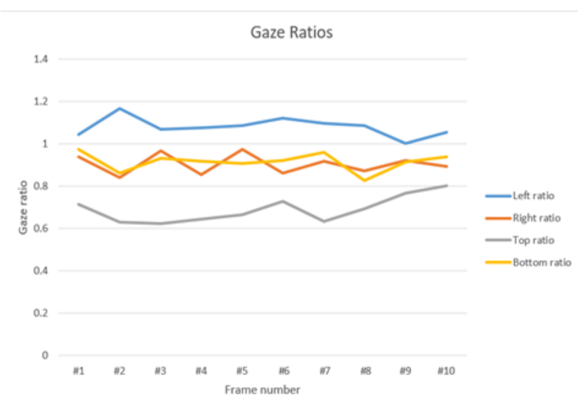

Fig. 4. Obtained gaze ratios for left, right, top and bottom directions

detect the pupil. This value varies drastically depending on the lighting conditions. The possibility of developing a solution that automatically sets the threshold value based on the environment can be explored. The application is intended to be capable of being used even by ALS patients who have completely lost all muscular capabilities, and only ocular movements remain. Thus it should provide for being operated completely by eye movements only. As a result, the application is designed to be activated when the patient blinks twice. This requires the camera to be continuously active, monitoring the eyes of the patient to detect a double blink. This leads to high power consumption and the possibility of unintended blink/gaze detections even when the patient is idle.

VII.FUTURE WORK

The possibility of developing the application for deployment as a mobile app on an Android/iOS platform is one that can be explored further. Additionally, the application can be extended to incorporate a notification system that sends messages to the caregiver's mobile phone when the patient is in need of something. For better power utilization and efficiency, the application can be made less dependent on eye movements and more touch-based controls can be provided for the caregiver.

ACKNOWLEDGMENT

We thank Dr. Ruby Abraham, Principal, Marian Engineering College for the help rendered towards the completion of this project and Dr. P Jayaprakash, Head of the Department of Computer Science and Engineering for providing us with all the facilities we required to carry out this work. We are also indebted to Ms. Sajni Nirmal, Assistant Professor, Department of Computer Science and Engineering, for her constant guidance and support.

REFERENCE

[1] Z. Aslam, A. Z. Junejo, A. Memon, A. Raza, J. Aslam and L. A. Thebo, "Optical Assistance for Motor Neuron Disease (MND) Patients Using Real-time Eye Tracking," 2019 8th International Conference on Information and Communication Technologies (ICICT), Karachi, Pakistan, 2019, pp. 61-65, doi: 10.1109/ICICT47744. 2019. 9001 922.

[2] Samantha C., Kristina D , Jeff D. and Carolyn E., Eye Gaze Technology, Nov. 5, 2012. Accessed

on: Jan. 23, 2021[Online]. Available: udlandat.weebly.com/eye-gaze-communication-technology.html

[3] Keiko Sakurai et al., ”Gaze Estimation Method Using Analysis of Electrooculogram Signals and Kinect Sensor,” Computational Intelligence and Neuroscience, vol. 2017, Article ID 2074752, 10 pages, 2017, DOI : 10.1155/2017/2074752.

[4] Anca Cehan et al., ”Keyword data base used in communication system with disabled people,” 2011 IEEE 17th International Symposium for Design and Technology in Electronic Packaging (SIITME), October 2011, DOI:10.1109/ SIITME. 2011.6102753.

[5] P. Viola and M. Jones, ”Rapid object detection using a boosted cascade of simple features,” Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[6] Davis, M., Popov, S. and Surlea, C. (2011). Real-Time Face Recognition from Surveillance Video. Studies in Computational Intelligence, 155–194. doi:10.1007/978-3-642-17554-1 8

[7] Anuradha Kar and Peter Corcoran, ”A Review and Analysis of Eye-Gaze Estimation Systems, Algorithms and Performance Evaluation Methods in Consumer Platforms”, IEEE Access, 2017, PP. 1-1. DOI:10.1109/ACCESS. 2017 .273 5633.

[8] T. Soukupova and J. Cech. (2016, Feb. 3) Real-Time Eye Blink Detection using Facial Landmarks. Center for Machine Perception, Department of Cybernetics Faculty of Electrical Engineering, Czech Technical University in Prague. Prague, Czech Republic. [Electronic]. Available: https://vision.fe. uni-lj.si/cvww2016 /proceedings /papers/05.pdf