Eye Pupil Tracking based App Option Selection for People with Cerebral Palsy

Sachin Shrestha
Assistant Professor
Electronics and Communication Engineering
Nepal Engineering College
Pokhara University
Bhaktapur, Nepal
sachins@nec.edu.np

Niraj Dhakal
Electronics and Communication Engineering
Nepal Engineering College
Pokhara University
Bhaktapur, Nepal
nearoz998@gmail.com

Pawan Dhakal
Electronics and Communication Engineering
Nepal Engineering College
Pokhara University
Bhaktapur, Nepal
pawan12430@gmail.com

Tek Raj Ojha
Electronics and Communication Engineering
Nepal Engineering College
Pokhara University
Bhaktapur, Nepal
tekrajojha6@gmail.com

***Abstract*— This paper is focused on the development of a system that could track eye pupils in order to select different options within a desktop application for children with cerebral palsy, having issues with verbal communication. In specific case of cerebral palsy called, "Spastic Quadriplegia/Quadriparesis", which is one of the most severe form of cerebral palsy, all four limbs, the trunk and the face are affected and the person have no control over his/her muscles limiting one's verbal communication. In such cases, they communicate by using some regular repetitive signs or gestures only with their parents, caretakers, and with persons who are regularly around them. In order to provide them the ability to communicate with people, this system with eye tracking functionality was developed. The system was built as a desktop application, using the python platform and had used a web-cam, monitor and an audio output of the laptop as additional components besides the processing unit. In the backend, the system performed image processing-based eye tracking algorithm, within the video captured from web-cam and generated audio output as per the selection made within the app. In the front-end part, the desktop app showed the window with image buttons that can be triggered with the input from the eye-movement and would generate audio output accordingly. The system showed the sensitivity of 86%, 76% and 80% for right movement of pupil, left movement of pupil and blink respectively. Likewise, specificity of 74%, 74% and 82%. The accuracies were obtained as 80%, 75% and 81% for each case respectively. The shifting speed was 614millisecond and the blink response detection was 1.023 seconds. The obtained results can be useful for the development of an enhanced voice communication system for people with cerebral palsy with verbal communication issues.**

Keywords*— *CP (Cerebral Palsy), Dlib, Eye Closure Ratio, Eye Gaze Ratio, Eye Gesture, Eye Monitoring, Eye Tracking, Histogram, OpenCV, Playsound, Pupils, Python, Spasticity (stiff muscles), Spastic Quadriplegia/Quadriparesis, UI (User Interface), Verbal Communication.

I. Introduction

A. Background and Problem Statement

Cerebral palsy (CP) is a group of disorders that affect a person's ability to move and maintain balance and posture caused by abnormal brain development or damage to the developing brain [1]. Around, 764,000 people are currently living with cerebral palsy, with 500,000 of them being children or teens and each year 8,000 to 10,000 infants are diagnosed with cerebral palsy [2]. According to Cerebral Palsy Alliance, one out of four children suffering from this disease is unable to talk [3].

Spastic cerebral palsy is the most common form, affecting up to 80% of people with CP resulting muscles to appear stiff and tight [4]. Among its several types, spastic quadriplegia/quadriparesis is the most severe form of spastic CP. This type affects all four limbs and often a person's torso, facial, and oral muscles are often affected, too. People with spastic quadriparesis usually cannot walk and often have other developmental disabilities such as intellectual disability; seizures; or problems with vision, hearing, or speech [5].

There has been a lot of research to develop systems that could aid in communication for people with CP [6] [7] [8]. Amongst them, the ones with eye tracking [7] can be more useful however all such systems are either expensive or with higher complexity. In case of Nepal, a very few researches have been made for development of high-tech devices targeting disability [9]. As such, there is a need to develop a system that could aid in verbal communication and at the same time is affordable for all.

B. *Objectives*

The major objective of this research paper is to develop a system that will have a python-based platform with visual front-end with image buttons that can be triggered with eye gestures and will consequently generate audio output. To achieve this, specific objectives are defined as:

- To develop a front-end with image buttons that will be triggered with eye-pupil gestures.
- To produce voice output based on eye gestures.

II. Related Works

A literature survey was carried on various technologies that are available for the people with cerebral palsy to analyze and to develop the suitable methodology required for the project. There are few technologies that contributed in communication with people with cerebral palsy. Few of the available techniques use Hough circle transformation for eye/pupil detection and point of gaze for screen coordinates [6] [10] [11]. Likewise various algorithms for frontal face detection followed by pupil/iris detection [12] and brain-computer-interface techniques [7] were also explored.

On the commercial side *SensoMotoric Instruments (SMI)* [13] and *Tobii* [14] are the main manufacturers of eye-tracking devices. While SMI focuses their production on eye tracker analysis systems, Tobii also produces some augmentative communication devices especially suitable to individuals with cerebral palsy. With these devices, people with disabilities can easily engage in daily communications via text or symbols, with word and phrase prediction to enhance the rate of communication.

III. Methodology

The high-level block diagram of the overall system is shown in Figure 1. The system consisted of a camera unit to capture the real-time video, a processing unit to compute the eye gesture tracking algorithm, a monitor unit to display the front-end and an audio unit to generate the audio as per the eye-gesture and the respective selection within the front-end.

The software section of the system was developed in python language platform, version 3.6.4.

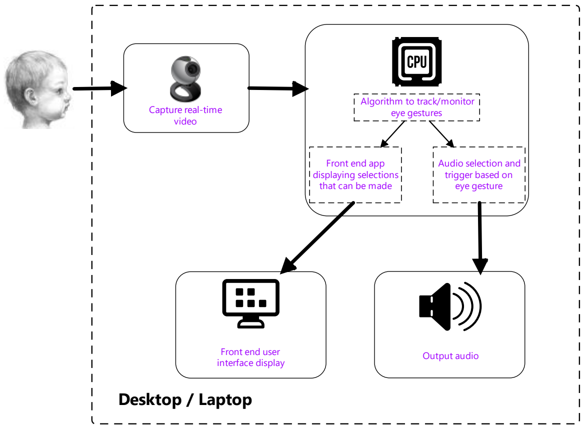

Figure 1 High Level Block Diagram of Overall System

The flowchart of the overall system is shown in Figure 2.

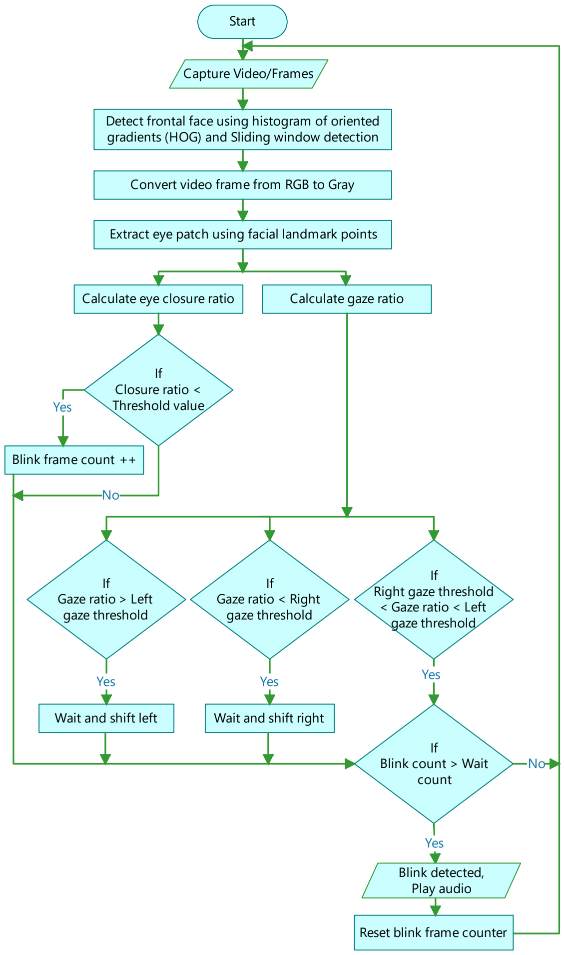

Figure 2 Overall Flowchart of the System

A. Capturing Real-Time Video

The camera of the PC was used for capturing real-time video. For the extraction and processing of images in real time, *OpenCV* library [15] within python platform was used.

B. Image Processing and Eye Monitoring

In order to detect faces, to extract eye pupils, and to define and detect different eye gestures, various levels of image processing were done.

1) Face Detection

Within the real-time video input, the face was detected using histogram of oriented gradient and sliding window detection. In image processing, an image histogram is a gray-scale value distribution showing the frequency of occurrence of each gray-level value of an image [16] [17] [18] [19]. A sliding window is a rectangular region of fixed width and height that slides across an image. On analyzing the histogram of the received image, the system was able to judge if the tonal distribution of a face within a frame of an input video corresponds to a face or not [20].

2) Eye Feature Extraction

Eyes were detected using image enhancement technique on the detected face. The input video frame containing a face was converted into a gray-scaled image. In this gray-scaled image, a library called *dlib* [21] portrayed 68 landmarks on the various positions of the detected face so as to represent the face completely using the coordinates of those landmarks. Knowing the numbering of each landmark points on the face, the eye patch was extracted from the face. This process of extracting eyes from the face is called as masking of the eyes. The eye pupils were detected with the image binarization technique.

3) Eye Closure Ratio

Eye closure ratio is the measure of eye, closed at the time of observation. The relevant flowchart is shown in figure 3. The eye closure ratio can be calculated as:

$$Eye\ Closure\ Ratio = \frac{Height\ of\ the\ open\ eye}{Width\ of\ the\ open\ eye}$$

4) Gaze Ratio

Gaze ratio was derived with the help of binarized image. It is the ratio of the white pixels in the left side of the eye to the white pixels in the right side of the eye. The relevant flowchart is shown in figure 3.

$$Gaze\ Ratio = \frac{No.of\ white\ pixels\ in\ vertical\ left\ of\ the\ eye}{No.of\ white\ pixels\ in\ vertical\ right\ of\ the\ eye}$$

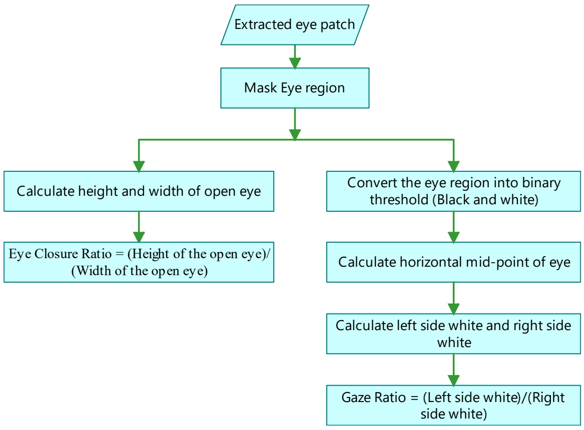

Figure 3 Flowchart of Eye Closure Ratio and Gaze Ratio Calculation

5) Eye Gestures

Eye gestures were defined for specific purposes to be used within the desktop app. In this app, three eye gestures were defined; looking left and looking right to navigate through the options in the app, and blinking of eyes to select the option. The right and left movement of the pupils were detected from the eye gaze ratio, according to the number of white pixels on either side of the vertical half of the eye patch. And the blink was detected from the eye closure ratio. In both of the cases, a threshold value was set for the detection.

C. Generating Voice Output

Following the detected eye gestures, a voice output was generated via a speaker when an option was selected by blinking. When an option was selected, a pre-recorded and stored audio file, with respect to the selected option was played. For example, when the user selected the 'banana' option, the audio file saying "I want a banana" was played via the speaker. This functionality was done with *playsound* [22], a python library to play the sound.

D. User Interface for User Interaction

The desktop app contained a user interface (UI) page as shown in Figure 4, with several options for selection. In order to navigate through the options, users had to move their eyes right or left according to the requirement and blink for selection of the option.

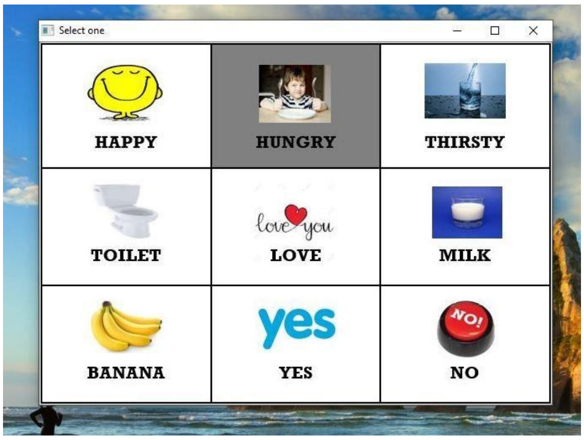

Figure 4 User Interface of App

IV. Result And Discussion

A. Evaluation Protocol

To measure the performance of the system outputs, parameters like true positive (TP), true negative (TN), false positive (FP) and false negative (FN) were calculated, based on the number of tests. In order to calculate those values, three major tests were performed with eye-pupil moved left, right and kept still. For all such three cases, 50 observations were made and above-mentioned parameters were calculated. The evaluation protocol to calculate those values were shown in Table 1.

Table 1 Evaluation Protocol to Calculate TP, FP, TN and FN

TP	Selection moves left and right with the corresponding left and right movement of the eye-pupil.
FP	Selection moves left and right even when the eye-pupil is in still position or, when eye-pupil is moved in the opposite direction i.e., right and left correspondingly.
TN	Selection remains unchanged when the eye-pupil is in still position.
FN	Selection remains unchanged even when the eye-pupil is moved left or right.

Based on the values of the TP, TN, FP and FN, various performance parameters were calculated such as Specificity (Sp), Sensitivity (Se) [23], Precision or Positive predictive value (PPV), Negative predictive value (NPV) and Accuracy (Acc). Next, considering a group with positive instances and negative instances of above conditions, the four outcomes were formulated into a 2×2 confusion matrix.

For the testing purpose, instead of using people with cerebral palsy, 3 male healthy persons were chosen of age 22.

B. Performance Evaluation of Eye Movement Tracking

With 50 different observations, the values for TP, FP, TN and FN were calculated. The observed values and the calculated values for different parameters are shown in Table 2.

Table 2 Summarized Result for Different Eye Gestures

Gesture	**Right**	**Left**	**Blink**
TP	43	38	40
FP	13	13	9
TN	37	37	10
FN	7	12	41
Specificity	0.74	0.74	0.82
Selectivity	0.86	0.76	0.80
PPV	0.7678	0.745	0.8136
NPV	0.8409	0.7551	0.8039
Accuracy	0.80	0.75	0.81

The confusion matrix for the different cases of eye-movement tracking based on 50 different observations is shown in Table 3.

Table 3 Confusion Matrix For Eye Movement

Actual Class / **Detected Class**	Right	Left	Still
Right	**43**	0	8
Left	0	**38**	5
Still	7	12	**37**

Next, the confusion matrix for different states of eye-blink detection based on 50 different observations is shown in Table 4.

Table 4 Confusion Matrix for Blink Detection

Detected Class \ **Actual Class**	Blinked	Not Blinked
Blinked	**40**	9
Not Blinked	10	**41**

From Table 2, it can be seen that the system detected right side movement of the pupils with 80% of accuracy, left side movement of the pupils with 75% accuracy, and the blink with 81% accuracy. The average accuracy of the system was calculated as 78.67%.

C. Performance Evaluation of Associated Delays

The time required by the system to shift left or right after the detection of gesture was calculated as 614 milliseconds. Also, the average blink response time was obtained as 1.023 seconds.

D. User's UI Interaction

The UI of the app was shown in Figure 5. The app UI was designed such that it would be simple to understand and easy to use. When the user moved their eyes right to navigate to the next option from that of the condition in Figure 4, the next consecutive option was highlighted as shown in Figure 5.

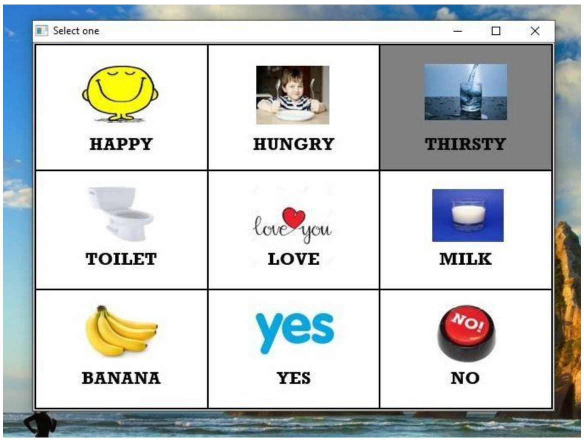

Figure 5 Next Option Highlighted When Looked Right

Next, when the user moved their eyes left to navigate to the previous option as in the position from Figure 4, the respective previous option was highlighted as shown in Figure 6.

That way, users could navigate through any option in the desktop app. In order to generate respective audio based on the selection made, users have to blink their eyes.

The system provides best performance if both the eyes of the user are normal. In case of squint condition of eyes, if at least one eye is healthy, the system can be modified to perform operations based on only one of the healthy eyes. In case if both of the eyes are squint, then the system can't work properly.

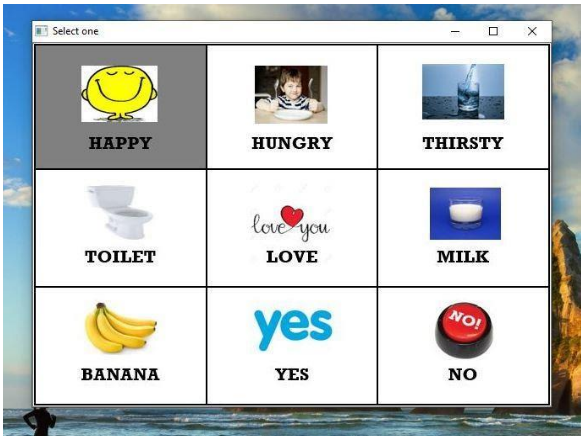

Figure 6 Previous Option Highlighted When Looked Left

Next, studies have shown that the human eye color varies from lightest blue to darkest of brown [24] [25]. In this model, since the grayscale conversion of the image of the eye patch has been done, the effect of the color of eye became insignificant. In the grayscale image of the eye patch, the intensity levels of the white part and iris were distinct. If the iris part of the eye at most had one third the intensity of the white part, the system could distinguish between the white part and the iris part and hence could detect the iris well.

Despite the fact that all tests were made with healthy persons as the main subject, it is believed that this experimentation could be the pilot project for the upcoming research on this topic and with the person with cerebral palsy as main subject for testing.

V. Conclusion And Future Development

This paper discussed the development of an eye pupil tracking based system that can identify different eye gestures and can generate different audio output based on the selections made within the application with eye-gesture. The system showed the sensitivity of 86%, 76% and 80% for right shift, left shift and blink respectively. Likewise, specificity of 74%, 74% and 82%. The accuracies were obtained as 80%, 75% and 81% respectively. The shifting speed was 614 millisecond and the blink response detection was 1.023 seconds. Thus, it can be concluded that the system can be a foundation for the enhanced voice communication system for people with cerebral palsy with verbal communication issues.

The developed system has limitations esp. with the restriction of its usage within the laptop, however can easily be ported into a more portable system consisting of raspberry pi. So as future developments, a more portable system can be created with smaller processing units and the tests within the subjects with cerebral palsy with issues with verbal communication.

ACKNOWLEDGEMENT

The authors would like to thank the team of specialists, educators and management from Self Help Group for Cerebral Palsy, for aiding the authors in getting to know the details of Cerebral palsy and the people affected with it. Also, the authors are equally grateful to Autism Care Nepal Society for their support with this study and allowing the authors to make a visit and letting the authors know more on disability. This paper is a part of the research work funded by University Grants Commission (UGC) Nepal and so authors are also equally thankful to UGC Nepal for their support.

REFERENCES

[1] "What is Cerebral Palsy?," [Online]. Available: https://www.cdc.gov/ncbddd/cp/facts.html. [Accessed October 2019].

[2] "Cerebral Palsy Prevalence and Incidence," [Online]. Available: https://www.cerebralpalsyguidance.com/cerebral-palsy/research/prevalence-and-incidence/. [Accessed October 2019].

[3] "How does cerebral palsy affect people?," [Online]. Available: https://cerebralpalsy.org.au/our-research/about-cerebral-palsy/what-is-cerebral-palsy/how-cerebral-palsy-affects-people/. [Accessed 06 October 2019].

[4] "Types of Cerebral Palsy," [Online]. Available: https://cparf.org/what-is-cerebral-palsy/types-of-cerebral-palsy/. [Accessed 05 October 2019].

[5] "Cerebral Palsy," [Online]. Available: https://www.cdc.gov/ncbddd/cp/facts.html. [Accessed October 2019].

[6] M. U. Ghani, S. Chaudhry, M. Sohail and M. N. Geelani, GazePointer: A Real Time Mouse Pointer Control implementation based on eye tracking, Lahore, 2013.

[7] Z. O. Abu-Faraj, M. J. Mashaalany, H. C. B. Sleiman, J.-L. D. Heneine and W. M. A. Katergi, "Design and Development of a Low-Cost Eye Tracking System for the completely locked-in patient," New York, 2006.

[8] Y. Guerrier, J. Naveteur, C. Kolski and F. Poirier, "Communication System for Persons with Cerebral Palsy," Cham, 2014.

[9] S. Shrestha and A. Shah, "Current Status of Assistive Technology for AAC for People," *SCITECH Nepal*, vol. 15, no. 1, pp. 36-44, December 2020.

[10] A. Galante and P. Menezes, "A gaze-based interaction system for people with cerebral palsy," Coimbra, 2012.

[11] C.-L. Jen, Y.-L. Chen, Y.-J. Lin, C.-H. Lee, A. Tsai and M.-T. Li, "Vision based wearable eye-gaze tracking system," Las Vegas, 2016.

[12] C. A. Perez, A. Palma, C. A. Holzmann and C. Pera, Face and eye tracking algorithm based on digital image processing, Santiago, 2001.

[13] "SensoMotoric Instruments," [Online]. Available: https://en.wikipedia.org/wiki/SensoMotoric_Instruments. [Accessed October 2019].

[14] "Products," Tobii Technology, [Online]. Available: https://tech.tobii.com/products/. [Accessed October 2019].

[15] "OpenCV," Intel, [Online]. Available: https://opencv.org/. [Accessed October 2019].

[16] M. A. ONABID and D. T. Charly, "Enhancing Gray Scale Images for Face Detection," *(IJACSA) International Journal of Advanced Computer Science and Applications,* vol. 8, 2017.

[17] V. Starovoitov, D. Samal and D. Briliuk, "Image Enhancement for Face Recognition," St. Petersburg, 2003.

[18] L. Tao, M.-J. Seow and V. K. Asari, "Nonlinear Image Enhancement to Improve Face Detection in Complex Lighting Environment," *International Journal of Computational Intelligence Research*, vol. 2, pp. 327-336.

[19] I. Bankman, "Processing and Analysis Management," in *Handbook of Medical Imaging*, Elsevier Inc., 2000, pp. 783-800.

[20] A. E. T. El_Deen, H. F. Hashem and A. Amin, "Application of Color Histogram to Human Face Detection," Egypt, 2008.

[21] "dlib C++ Library," [Online]. Available: http://dlib.net/. [Accessed October 2019].

[22] "Playsound," [Online]. Available: https://pypi.org/project/playsound/. [Accessed October 2019].

[23] J. Yerushalmy, "Statistical Problems in Assessing Methods of Medical Diagnosis, with Special Reference to X-Ray Techniques," *Public Health Reports,* vol. 62 (40), pp. 1432-1449, 3 October 1947.

[24] R. Sturm and M. Larsson, "Genetics of human iris color and patterns," *Pigment cell & melanoma research,* vol. 22, pp. 544-62, 1 August 2009.

[25] W. A. H. P. e. a. Liu F, "Digital Quantification of Human Eye Color Highlights Genetic Association of Three New Loci," *PLOS Genetics*, vol. 6, pp. 1-15, May 2010.