Vol. 2 - No. 2
March 2025 E-ISSN: 2584-0657

Enhancing Communication for Quadriplegic Patients through Gaze Controlled Voice Messaging

Eindhumathy J[1], Thenmozhi M[2], Sana Azeemah Y[3], Sharvina V[4]
[1]Assistant Professor, Department of Electronics and Communication Engineering
Saranathan College of Engineering, Tiruchirappalli, India
[2,3,4]Student, Department of Electronics and Communication Engineering
Saranathan College of Engineering, Tiruchirappalli, India
DOI: *https://doi.org/10.5281/zenodo.15704609*

Abstract
Quadriplegic patients face communication barriers due to limited mobility, requiring innovative solutions. This project introduces a gaze-controlled voice messaging system using MediaPipe's eye tracking and face mesh technology for real-time, high-precision gaze detection. Arduino manages the hardware components, while Python handles the software logic, ensuring seamless integration and responsiveness. The system features a graphical interface where users can select messages through gaze direction. Selected messages are converted into speech using text-to-speech and transmitted via Bluetooth to caregivers. An LCD display provides real-time confirmation of the chosen message, minimizing errors and enhancing user confidence. One of the key strengths of our system is its efficiency—the gaze detection operates with minimal latency, ensuring real-time interaction. The modular architecture and lightweight processing requirements make it suitable for low-power devices, enhancing usability in both home and clinical environments. Compared to existing systems, which are often expensive, require specialized hardware, or rely heavily on cloud processing, our solution is cost-effective, portable, and entirely offline- capable. It uses open source technologies and widely available components, making it accessible for patients with limited financial resources. Our project stands out due to it slow cost with high accuracy, real-time feedback loop via LCD, wireless communication through Bluetooth, minimal learning curve and user-friendly design, and easy scalability and upgradability using MediaPipe's modular nature. This solution empowers quadriplegic patients with independent, reliable, and efficient communication, combining the accuracy of MediaPipe, the flexibility of Python, and the control of Arduino to create a robust assistive technology for daily living.
Keywords*: Gaze-controlled communication, MediaPipe Face Mesh, Quadriplegic assistance, Arduino integration, Python-based system, Text-to-speech (TTS)*

INTRODUCTION

Quadriplegic patients face significant communication challenges due to the loss of motor functions. Traditional systems often have complex interfaces, making communication difficult. This project proposes a gaze-controlled voice messaging system using MediaPipe's face mesh and eye-tracking modules for real time gaze detection. By tracking eye movements, the system enables patients to interact with a Graphical User Interface (GUI) and select messages effortlessly.

The system integrates Arduino for hardware control and Python for software processing. Arduino manages hardware interactions, while Python translates gaze inputs into GUI actions. Bluetooth connectivity allows wireless transmission of selected messages to a mobile application, where they are played as voice output, ensuring effective caregiver communication.

An LCD display provides real-time feedback, allowing users to confirm their choices before transmitting messages, reducing errors and improving efficiency. Designed to be cost-effective and user-friendly, the system leverages accessible technologies like Arduino, Python, and MediaPipe, making it a viable alternative to expensive existing solutions.

With a modular and customizable design, the system can be enhanced with features like predictive text and personalized messages. By integrating gaze detection, hardware control, and voice playback, this project empowers quadriplegic patients with an intuitive communication tool, improving their independence and quality of life.

BACKGROUND AND PREVIOUS WORK

In [1], an uncertainty-aware gaze tracking system is introduced for use in assisted living environments. Traditional gaze tracking methods often struggle with challenges such as lighting variations, head movements, and user-specific differences, which can lead to reduced accuracy. To address these limitations, the system in [1] models the uncertainty in gaze measurements, enhancing its robustness and reliability in real-world conditions.

The work in [5] presents a non-calibrated eye-tracking system aimed at augmenting stroke diagnosis.

This system offers a cost-effective and rapid solution, especially valuable in emergency medical settings where time is critical. Unlike traditional stroke diagnosis methods, the approach in [5] allows for quick assessments by analyzing eye movement features such as saccades, fixations, and smooth pursuit. The non-calibrated nature of the tracker simplifies clinical deployment, while machine learning techniques are employed to distinguish between normal and abnormal patterns. This aids healthcare providers in making faster and more informed decisions.

In [1], a system is developed to assist individuals with speech impairments in communication by using eye blinks. Known as Blink-to-Live, this system uses a high-resolution camera and a robust algorithm that can effectively differentiate between voluntary and involuntary blinks, allowing users to convey messages with greater clarity and ease.

Finally, the study in [2] introduces a 3D gaze estimation system designed for head-mounted eye trackers. One of the key innovations of this work is an auto-calibration method that significantly reduces calibration errors caused by head movements. The system utilizes high-resolution infrared cameras to monitor pupil positions and corneal reflections, and leverages a machine learning-based approach to enable self calibration, improving the system's adaptability and accuracy.

METHODOLOGY

Eye Movement Detection

Eye movement detection is fundamental to gaze-controlled systems, enabling users to interact through their eye motions alone. It involves tracking eye positions and movements—such as saccades, fixations, and smooth pursuits—using infrared sens sensors, high-resolution cameras, and computer vision algorithms. These movements are analyzed to determine user intent and link gaze points to interface actions. MediaPipe enhances this process with machine learning-based facial landmark detection for real-time gaze estimation. In this project, MediaPipe’s modules help quadriplegic patients select and transmit messages via a digital interface, offering a non-invasive, intuitiveway to communicate with caregivers efficiently and accurately.

MediaPipe

MediaPipe is an open-source framework by Google that supports affordability, ease of use, and adaptability. Its modular design allows future enhancements like voice control and predictive text input. By integrating MediaPipe’s gaze detection, Arduino’s hardware control, and Python’s software logic, it offers a reliable communication platform, empowering quadriplegic patients with greater independence.

PROPOSED METHOD

The gaze-controlled voice messaging system enhances assistive communication for quadriplegic patients by integrating computer real-time computer vision tasks like hand, face, and eye tracking. It operates efficiently on various platforms, including mobile and embedded systems, thanks to its modular, graph-based architecture. This structure allows parallel data processing with minimal latency, making it ideal for real-time gaze tracking. In this project, MediaPipe’s face and eye-tracking modules are used to detect the user's gaze and translate it into interface selections. These capabilities allow quadriplegic patients to communicate effectively through a visual interface, ensuring high accuracy, smooth system interaction, and improved quality of life.

Face Mesh

MediaPipe’s face mesh module tracks 468 facial landmarks to construct a 3D facial map in real time. It accurately detects features like eyes, nose, and mouth using deep learning models, even under varying lighting and angles. This high-fidelity mesh enables precise gaze tracking by capturing subtle facial movements and eye positions. In the proposed system, face mesh is crucial for detecting gaze direction, which controls interface actions. This allows quadriplegic users to interact hands-free with minimal effort. The module's precision and responsiveness make it ideal for enabling intuitive communication, significantly enhancing user independence and accessibility.

EXISTING SYSTEM

Quadriplegic patients face severe communication challenges due to motor function loss in all four limbs and the torso. Traditional voice-based communication systems often fail due to complexity and limited adaptability. This project proposes a gaze-controlled voice messaging system using MediaPipe’s face mesh and eye tracking modules for efficient communication.

MediaPipe, an open-source framework by Google, enablesreal-time gaze detection and facial tracking. The system translates eye movements into commands, allowing patients to interact with a graphical user interface (GUI). Arduino manages hardware components like an LCD, while Python processes gaze data.

Bluetooth connectivity enables wireless communication, transmitting messages to caregivers via a mobile app with voice playback. The LCD provides real-time feedback, reducing errors. This cost-effective system uses widely available technologies, making it more accessible than expensive alternatives. Unlike existing solutions, this system prioritizes vision, hardware-software interaction, and wireless connectivity. Using MediaPipe's face mesh and eye-tracking modules, it detects gaze direction and translates it into commands for selecting predefined messages on a graphical interface.

An Arduino microcontroller manages hardware components, including an LCD and Bluetooth module, while Python processes gaze data in real time. The LCD provides feedback, allowing users to confirm selections before sending messages, minimizing errors. Once confirmed, messages are transmitted via Bluetooth to a mobile application using the Dabble app, where they are converted into voice output via text-to-speech (TTS), ensuring prompt caregiver communication. The interface reduces eye strain with adjustable dwell time settings.

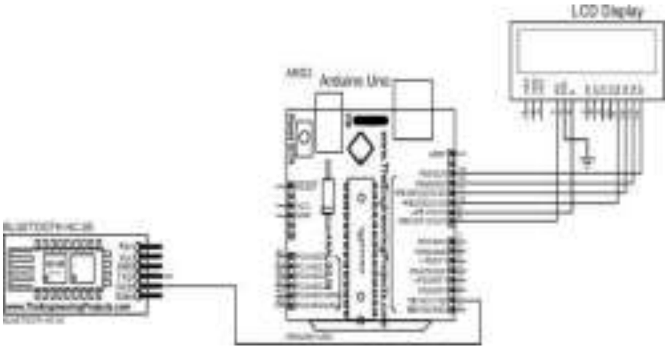
Error-checking algorithms enhance accuracy by adapting to lighting and facial variations, ensuring reliable performance. A built-in emergency alert system notifies caregivers in urgent situations.

- Hands-free communication using gaze tracking.
- Real-time feedback through an LCD, reducing errors.
- Bluetooth-enabled wireless message transmission.
- High-precision gaze detection with MediaPipe.
- Cost-effective, customizable, and scalable.
- Adaptive dwell time settings to reduce fatigue.
- Reliable error-checking mechanisms.

By integrating high-precision gaze detection, Arduino-based hardware, Python-driven processing, and Bluetooth connectivity, the system enhances independence and quality of life for quadriplegic patients.

Block Diagram

Circuit Diagram



Future enhancements could include:

- Custom message input for personalized communication.
- Multi-language support for broader accessibility.

Combining Python, MediaPipe, and Arduino, the system is cost effective, user friendly, and efficient. It significantly improves the independence and quality of life of quadriplegic patients by enabling effective communication with caregivers healthcare providers, and loved ones. Designed for real-world use, the system

IMPLEMENTATION

The Python-based gaze-controlled voice messaging system is designed to assist quadriplegic patients by enabling hands-free communication through real-time eye tracking. Utilizing MediaPipe's Face Mesh module, the system captures live video input via a webcam and detects facial landmarks to interpret gaze direction accurately. Users can navigate a set of predefined voice messages displayed on a graphical user interface (GUI) using subtle eye movements, enabling intuitive and accessible message selection.

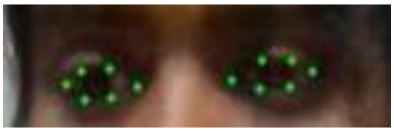
Once a message is selected, it is sent via Bluetooth to an external device using PySerial for communication with an Arduino module. This ensures reliable hardware interaction, such as triggering an external speaker. A text-to-speech (TTS) engine converts the selected message into audible speech, allowing caregivers to hear it clearly. An LCD screen provides real-time feedback, confirming the user's selection before transmission, reducing errors and improving usability. By integrating computer vision, Bluetooth communication, and TTS technology, the system offers a practical and efficient communication solution for non verbal users.

It ensures smooth interaction between users and caregivers, enhancing independence for individuals with severe motor impairments. The system is optimized for accuracy, low latency, and user-friendliness in real-time conditions. accounts for variations in lighting, head movement, and user specific differences.

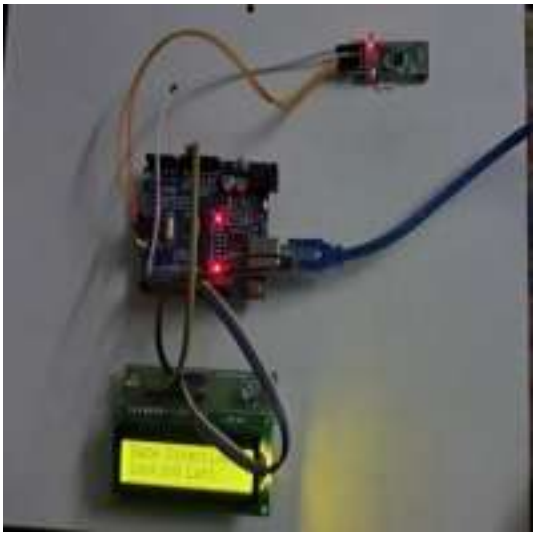
Python Code Snippet

OUTPUT

The implemented gaze tracking system effectively detects and interprets a user's eye movements in real time to determine gaze direction. Using computer vision algorithms—such as MediaPipe Face Mesh or deep learning-based CNNs—the system captures frames from a webcam and highlights key points around the eyes with green markers.



These markers help analyze the orientation of the eyeballs to classify gaze directions, such as left, right, up, down, or straight. The Python script processes these movements and translates them into predefined commands, which are sent via serial communication to an Arduino microcontroller.



The system is built on open-source platforms like MediaPipe and Arduino, keeping it cost-effective and accessible to users with limited financial resources. Its modular design supports future upgrades and feature expansion.

Designed for comfort and usability, the system features adjustable dwell time settings to meet varying user needs and reduce fatigue.

Moreover, the system showcases the powerful synergy between computer vision, machine learning, and embedded systems. It performs efficient edge computing on low-power devices, eliminating reliance on cloud infrastructure and reducing latency. The modular design allows it to be adapted for various applications, including home automation, wheelchair control, or even gaming and security.

In summary, the gaze tracking system successfully bridges human intention and machine action, offering real-time, hands-free control with wide-ranging applications in assistive technology and beyond.

CONCLUSION

The proposed gaze-controlled voice messaging system offers an innovative and accessible communication solution for quadriplegic patients. Using MediaPipe's face and eye-tracking modules, it enables accurate real-time gaze detection, allowing users to interact with a graphical interface by focusing on message options. Patients can select, confirm, and transmit messages via a Bluetooth-enabled mobile application.

Python handles the software logic, while Arduino manages hardware control, ensuring seamless integration. Real-time feedback is provided through an LCD display, allowing users to verify selections and minimize miscommunication.

Bluetooth connectivity supports wireless voice playback, enhancing portability and ease of use in different environments An intuitive interface minimizes cognitive load, making operation simple and user-friendly. Error-handling mechanisms and adaptive algorithms contribute to reliable performance across diverse conditions.

By combining real-time gaze tracking, Bluetooth communication, and text-to-speech, the system significantly enhances the quality of life for quadriplegic individuals through a scalable and customizable assistive technology solution.

REFERENCES

[1] M. Ezzat, M. Maged, Y. Gamal, M. Adel, M. Alrahmawy, and S. El-Metwally, ‘‘Blink-to-live eye-based communication system for users with speech impairments,’’ Sci. Rep., vol. 13, no. 1, p. 7961, May 2023, doi: 10.1038/s41598-023-34310-9.

[2] M. Liu, Y. Li, and H. Liu, “3D gaze estimation for head mounted eye tracking system with auto-calibration method,” IEEE Access, vol. 8, pp. 104207–104215, 2020, doi: 10.1109/ACCESS.2020.2999633.

[3] H. Zhang, S. Wu, W. Chen, Z. Gao, and Z. Wan, “Self calibrating gaze estimation with optical axes projection for head-mounted eye tracking,” IEEE Trans. Ind. Informat., vol. 20, no. 2, pp. 1397–1407, Feb. 2024, doi: 10.1109/TII.2023.3276322.

[4] Z. Hu, S. Li, C. Zhang, K. Yi, G. Wang, and D. Manocha, “DGaze: CNN-based gaze prediction in dynamic scenes,’’ IEEE Trans. Vis. Comput. Graphics, vol. 26, no. 5, pp. 1902–1911, May 2020, doi: 10.1109/TVCG.2020.2973473.

[5] M. A. Hassan, C. M. Aldridge, Y. Zhuang, X. Yin, T. McMurry, G. K. Rohde, and A. M. Southerland, ‘‘Approach to quantify eye movements to augment stroke diagnosis with a non-calibrated eye-tracker,’’ IEEE Trans. Biomed. Eng., vol. 70, no. 6, pp. 1750–1757, Jun. 2023, doi: 10.1109/TBME.2022.3227015.

[6] M. Yang, C. Cai, and B. Hu, “Clustering based on eye tracking data for depression recognition,’’ IEEE Trans. Cognit. Develop. Syst., vol. 15, no. 4, pp. 1754–1764, Dec. 2023, doi: 10.1109/TCDS.2022.3223128.

[7] V. Yaneva, L. A. Ha, S. Eraslan, Y. Yesilada, and R. Mitkov, “Detecting high-functioning autism in adults using eye tracking and machine learning,’’ IEEE Trans. Neural Syst. Rehabil. Eng., vol. 28, no. 6, pp. 1254–1261, Jun. 2020, doi: 10.1109/TNSRE.2020.2991675.

[8] P. Kowalczyk and D. Sawicki, “Blink and wink detection as a control tool in multimodal interaction,’’ Multimedia Tools Appl., vol. 78, no. 10, pp. 13749–13765, May 2019, doi: 10.1007/s11042-018-6554-8.