International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Cursor Control and Speech Transcription for Paralyzed Individuals

K. Akhil Abhilash [a]*, K. Vamsi [a], K. Anirudh Trivedi [a], M. Dathatreya [a], Y. Sai Suresh [a]

[a] *Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh, India.*

ABSTRACT

Computers are an essential part of modern life, but individuals with disabilities, particularly those who are paralyzed, may encounter difficulties in using them. Paralysis can severely limit movement to just the eyes, head, and voice, making computer use challenging without assistance. Although several techniques have been developed to improve cursor control systems, such as voice and gesture control, some users have found these methods inconvenient or challenging to use. To address these issues, a new model has been proposed that eliminates the need for assistance. This model identifies specific facial landmarks that allow for the control of cursor movement and clicking events without requiring direct physical contact with the user. The system utilizes eye gaze and facial movements to achieve this, and it also incorporates an Integrated Speech-to-Text system for textual interpretation. The system has a user-friendly interface that makes it easy to use and provides an improved user experience.

Keywords: Shape Predictor, OpenCV, Human Computer Interface (HCI), Speech Recognition, Pyaudio

1. Introduction

Paralysis can have a significant impact on a person's life, as it often results in the loss of control or sensation in their hands and legs. With the increasing reliance on computers in various aspects of life, it can be challenging for paralyzed individuals to use digital devices independently. They may need to rely on others for assistance, which can be inconvenient and limit their independence. To address this issue, a proposed system aims to create a cursor control model and a speech-to-text model that would enable paralyzed individuals to use computers independently. The system uses facial landmarks as input for the cursor control system. During the calibration process, the system creates a vector format using points from the face to control the cursor's movement based on the user's facial movements. The system also configures mouse clicking events during the calibration process, using left and right eye blinking and a wide-open mouth for mouse scrolling. The open-cv package is used to detect the necessary features from the face, and the Gaussian filter is applied to plot a graph to check the probability of the cursor clicking and movement events. In addition to the cursor control model, the proposed system includes a speech-to-text model that uses Google API to detect speech from the user and convert it into recognized text. The listening duration for the speech can be customized based on the user's needs. Overall, the proposed system has the potential to significantly improve the quality of life for paralyzed individuals by providing them with greater independence in using digital devices.

Cursor Control and Speech Transcription system for Paralyzed Individuals is an innovative technology developed to assist individuals who have paralysis in operating their computers through voice commands. This system enables users to control the cursor on their computer screen, type text, and navigate through various applications without the need for hand movements. It provides a means for individuals with limited mobility to access the internet, communicate with others, and perform various tasks with ease. The system functions by converting the user's speech into text, which is then used to manipulate the computer's cursor and carry out other functions. This technology has the potential to enhance the quality of life for people with paralysis by enabling them to stay connected with the world around them.

2. Related Work

[1]. The authors have proposed a novel model for Human-Computer Interaction (HCI) that facilitates communication with the system and enhances usability. This model aims to develop a Human-Machine Interaction (HMI) system that provides assistive communication technology for individuals with disabilities and offers greater flexibility. Traditional computer interfaces, which rely on standard mice, require users to have a hand and a flat surface for cursor movement. To overcome these limitations, the concept of an "air mouse" was introduced, which allows users to control the mouse pointer and perform clicking events using hand gestures and eye movements. The location of head gestures is determined using a three-axis accelerometer and gyroscope combination.

[2]. The paper introduces an innovative approach to designing a human-computer interface (HCI) that is specifically tailored to individuals suffering from neurodegenerative conditions, such as ALS and MS. The proposed system involves using eye-tracking devices to capture eye movements, which are then

analyzed using deep learning algorithms to identify the user's intention and operate the computer interface accordingly. The study's findings demonstrate that this new approach significantly outperforms existing systems in terms of accuracy, speed, and usability, providing an effective solution for individuals with neurodegenerative conditions to communicate and interact with computers.

[3]. The proposed method involves utilizing a webcam to capture images of the user's eyes. Subsequently, the images are processed to extract various features, including the pupil's location and the direction of gaze. These features are then analyzed using a set of algorithms to determine the user's intent, allowing them to perform actions like moving the cursor or clicking on buttons. This technique has proven to be successful in helping people with disabilities use a computer without relying on conventional input devices.

[4]. This paper describes a touch-free computer navigation system that uses human eye movements as a control mechanism. The purpose of this system is to provide an alternative means of computer input for individuals who are unable to use conventional devices like a mouse or keyboard. The system works by utilizing an electro-oculogram (EOG) device to track eye movements and convert them into mouse pointer movements. The experimental results demonstrate that the system is highly accurate and reliable in providing computer control without the need for physical input devices. This technology has the potential to greatly improve the accessibility of computer usage for individuals with physical disabilities.

[5]. The paper introduces a novel cursor control system designed specifically for individuals who suffer from motor disabilities. The system is based on the electro-oculography (EOG) signals generated by eye movements and is intended to serve as a low-cost and accessible alternative to traditional mouse and keyboard interfaces. The authors demonstrate the system's feasibility and its potential to aid disabled individuals in performing daily computer tasks.

[6]. The paper presents a novel approach to improve computer accessibility through a vision-based interface. The system enables users to operate the computer using head movements, which are tracked by a low-cost camera and translated into cursor commands. The system's efficacy in enabling computer use for individuals with physical disabilities was demonstrated through tests. The system eliminates the need for physical contact, making it a cost-effective and efficient alternative.

[7]. The authors have presented a novel system for individuals with disabilities that allows them to control the computer through head movements. The system employs a camera to track the user's head movements and converts them into mouse movements, eliminating the need for manual control. The results of the study demonstrated that the system was effective in terms of accuracy, efficiency, and user satisfaction, indicating that it has the potential to improve the quality of life for people with disabilities.

[8]. The article introduces a novel system that enables users to control the computer cursor through eye and face movements. The system leverages both facial landmarks and eye tracking to recognize and interpret user gestures. It comes with a graphical user interface that provides instant feedback on the identified movements, and can serve as an assistive technology for individuals with disabilities or as a fresh way of interacting with computers.

[9]. The proposed system utilizes deep learning to detect and track eyeball movement for controlling the cursor. It employs a Convolutional Neural Network (CNN) to extract features and classify the movement of the eyeball accurately. According to the research paper, the system was able to achieve a high accuracy of 75-80% in detecting and classifying the eyeball movements for cursor control.

[10]. This paper introduces a system for converting speech to text using neural network methods based on deep learning. The system incorporates the Mel-frequency cepstral coefficients technique for feature extraction and utilizes the long short-term memory model for the conversion of the speech signal into text. Through the proposed method, the system was able to achieve high levels of accuracy in accurately transcribing speech into text.

[11]. The paper introduces a novel technique for speech recognition that utilizes both audio and visual modalities to enhance the recognition performance. The authors propose a multimodal recurrent neural network model that leverages the complementary information from both modalities. According to the findings, the proposed method surpasses the conventional single-modality-based approaches in terms of recognition accuracy. The method involves using deep learning-based assistive technology to improve speech recognition accuracy by incorporating both audio and visual data. The paper provides a comprehensive overview of the system's architecture, including the integration of convolutional neural networks and long short-term memory networks. The study also includes experimental results that highlight the effectiveness of the proposed system in enhancing speech recognition accuracy for individuals with hearing impairments.

[12]. The paper describes an assistive technology that utilizes deep learning to enhance audio-visual speech recognition for individuals who have hearing impairments. This technology employs a multimodal technique that integrates both audio and visual cues. The system was able to achieve a recognition accuracy of 95%, surpassing current technologies.

[13]. The authors have put forward a new system for controlling a cursor without the need for hands, which relies on natural movements of the head and subtle twitches of the cheek muscles. To accomplish this, the system employs a webcam to monitor the user's head movements and an EMG sensor to detect the cheek muscle twitches, thereby creating an interface that is more user-friendly and accessible.

[14]. The proposed system utilizes head pose detection to enable users to control the mouse cursor with their eye movements, providing a more efficient and user-friendly way for people with physical disabilities to operate their computers without requiring manual dexterity. Pupil Mouse is a computer interaction system that enables users to operate their computers hands-free using eye gaze and head pose detection. The system employs a camera mounted on the user's head to capture their head pose and track their eye movements, which are then translated into movements of the mouse cursor on the screen. Pupil Mouse offers an alternative to traditional mouse and keyboard input methods, which can be challenging for individuals with motor disabilities, by providing a more natural and intuitive means of interacting with a computer. The technology has the potential to enhance the quality of life for people with motor disabilities and facilitate their participation in daily activities.

[15]. The researchers created a communication system that utilizes electrooculography (EOG) and can be controlled through eight-directional eye movements. This system is an affordable and non-invasive solution for individuals with physical disabilities, and the tests conducted showed a high level of precision in controlling the system.

[16]. The proposed system, called the camera mouse, allows individuals with severe disabilities to control a computer cursor using their head movements. This is achieved through the use of a video camera and specialized computer vision algorithms. The camera mouse system provides an accessible means for disabled individuals to interact with various devices.

[17]. In this paper, we introduce a novel human-computer interface (HCI) utilizing electrooculogram (EOG) signals and deep learning techniques. Our system detects and interprets eye movements via EOG signal acquisition, pre-processing, feature extraction, and deep learning-based classification. Specifically, we employ a convolutional neural network (CNN) to accurately classify various eye movements. Our proposed HCI has the potential to improve accessibility for individuals with disabilities, especially those with limited mobility or communication impairments. We evaluate the effectiveness of our system using a dataset of EOG signals recorded from multiple participants. The results show that our HCI is capable of real-time control of various computer applications through accurate classification of eye movements.

[18]. The paper introduces a novel approach to codify electrooculogram (EOG)-based eye movements for human-computer interaction. Eye movements play a crucial role in human-computer interaction, enabling users to interact with graphical user interfaces and virtual environments. The proposed method leverages EOG signals to detect various types of eye movements, such as saccades, smooth pursuits, and fixations. The signals undergo processing using a custom algorithm that extracts features like amplitude, duration, and velocity. These features are then used to generate a code that represents the specific eye movement type. The effectiveness of the method is evaluated using a dataset of EOG signals recorded from several subjects, demonstrating high accuracy in detecting and classifying various eye movements. This method has the potential to enhance the precision and speed of eye tracking systems, making them more suitable for use in real-world applications, including virtual reality, gaming, and assistive technologies.

[19]. Eye tracking technology has made significant advancements in recent years, providing accurate detection of where individuals look on a computer screen. This article examines the use of eye movements as a human-computer interaction interface, with various techniques developed to create more natural and intuitive interactions. The article explores the advantages and challenges associated with using eye movements as the primary input method, as well as its potential applications in gaming, assistive technology, and virtual reality. Overall, this technology has the potential to improve the user experience.

[20]. The Eye-blink detection system for human-computer interaction is a technology that can improve the user's interaction with computers by detecting eye-blinks. This system utilizes the analysis of the electrical signals produced by the eye muscles during blinking and has various applications, including video gaming, virtual reality, and hands-free computing. The system can detect eye-blinks and use them to control different functions of a computer or device, such as scrolling, zooming, and selecting items on a screen. Overall, this technology has the potential to enhance the user experience and provide new possibilities for human-computer interaction.

[21]. This paper describes a system that enables cursor control through eye movements. The system incorporates an eye tracking device to detect and interpret the user's eye movements, which are then translated into corresponding movements of the computer cursor on the display. The proposed system offers a potential alternative for individuals with disabilities or those who find conventional input devices such as a mouse or touchpad challenging to use. The system is designed to be user-friendly and intuitive, requiring minimal training or calibration. Experimental results indicate the feasibility and efficacy of the system in providing precise and responsive cursor control through eye movements.

[22]. The paper introduces a novel approach for controlling a computer cursor using real-time 3D face tracking technology. The system employs a 3D camera to track the user's facial movements and translate them into corresponding cursor movements on the screen. The advantage of this hands-free system is that it eliminates the need for a mouse or other input device, making it more accessible for people with disabilities and providing a more intuitive user interface. The abstract emphasizes the benefits of this technology for improving accessibility and enhancing the user experience.

[23]. This paper presents a novel system that utilizes machine learning to enable individuals to control a computer cursor using their eye movements. The system comprises an eye tracking device, machine learning algorithms, and a computer with a graphical user interface. By capturing eye movements, the eye tracking device provides input to the machine learning algorithms that interpret the movements and translate them into intended cursor movements. This system is designed to assist individuals with physical disabilities or injuries that prevent them from using conventional input devices such as a mouse or keyboard. A series of tests are conducted to evaluate the accuracy and effectiveness of the system, demonstrating its potential as an assistive technology for individuals with limited mobility.

3. Conclusion

After successfully implementing the models with a unified interface, users can conveniently switch between them based on their preference. The models demonstrated excellent performance in both cursor control and speech-to-text conversion processes. Consequently, this project offers people with paralysis the ability to work independently on computers and eliminates the need to rely on others for personal or professional tasks. The facial landmark-based cursor control and speech-to-text system has demonstrated great potential in significantly improving the quality of life for individuals with paralysis. By utilizing facial landmarks to control a computer cursor and speech-to-text technology for communication, this system allows for a level of independence and communication that was previously unattainable for those living with paralysis. Several studies have shown that this technology has a high level of accuracy and usability, but further research is necessary to optimize it for different types of paralysis and improve its reliability. In conclusion, the facial landmark-based cursor control and speech-to-text system is a valuable tool for individuals with paralysis, and has the potential to revolutionize the way they interact with the world around them. Further development and research in this field could lead to significant advancements in assistive technology and greatly improve the quality of life for those living with paralysis.

References

[1]. Fahim, S. R., Datta, D., Sheikh, M. R. I., Dey, S., Sarker, Y., Sarker, S. K., ... & Das, S. K. (2020). A visual analytic in deep learning approach to eye movement for human-machine interaction based on inertia measurement. IEEE Access, 8, 45924- 45937.

[2]. Ramakrishnan, J., Doss, R., Palaniswamy, T., Faqihi, R., Fathima, D., & Srinivasan, K. (2022). High performance computation of human computer interface for neurodegenerative individuals using eye movements and deep learning technique. The Journal of Supercomputing, 1-21.

[3]. Deepika, S. S., & Murugesan, G. (2015, March). A novel approach for Human Computer Interface based on eye movements for disabled people. In 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) (pp. 1-3). IEEE.

[4]. Lewandowski, T., & Augustyniak, P. (2010, May). The system of a touchfree personal computer navigation by using the information on the human eye movements. In 3rd International Conference on Human System Interaction (pp. 674-677). IEEE.

[5]. Eduardo, P., & Xavier, M. (2014, June). Cursor control system of a computer by electrooculographs signs for motor disability. In 2014 IEEE Canada International Humanitarian Technology Conference-(IHTC) (pp. 1-4). IEEE.

[6]. Varona, J., Manresa-Yee, C., & Perales, F. J. (2008). Hands-free vision-based interface for computer accessibility. Journal of Network and Computer Applications, 31(4), 357-374.

[7]. Abiyev, R. H., & Arslan, M. (2020). Head mouse control system for people with disabilities. Expert Systems, 37(1), e12398.

[8]. Dongre, A., Pinto, R., Patkar, A., & Lopes, M. (2020, July). Computer Cursor Control Using Eye and Face Gestures. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.

[9]. Sharanyaa, S., & RP, M. (2021, July). Eyeball Cursor Movement Detection Using Deep Learning. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC).

[10]. BABU PANDIPATI, D. R. (2021). Speech to text Conversion using Deep Learning Neural Net Methods. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(5), 2037-2042.

[11]. Feng, W., Guan, N., Li, Y., Zhang, X., & Luo, Z. (2017, May). Audio visual speech recognition with multimodal recurrent neural networks. In 2017 International Joint Conference on neural networks (IJCNN) (pp. 681-688). IEEE.

[12]. Kumar, L. A., Renuka, D. K., Rose, S. L., & Wartana, I. M. (2022). Deep learning based assistive technology on audio visual speech recognition for hearing impaired. International Journal of Cognitive Computing in Engineering, 3, 24-30

[13]. Kumar, L. A., Renuka, D. K., Rose, S. L., & Wartana, I. M. (2022). Deep learning based assistive technology on audio visual speech recognition for hearing impaired. International Journal of Cognitive Computing in Engineering, 3, 24-30.

[14]. Iwata, M., & Ebisawa, Y. (2008, July). PupilMouse supported by head pose detection. In 2008 IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems (pp. 178-183). IEEE.

[15]. Cao, Y., Liu, S., Peng, Y., & Li, J. (2020). DenseUNet: densely connected UNet for electron microscopy image segmentation. IET Image Processing, 14(12), 2682-2689.

[16]. Hu, G., Hu, Z., Liu, J., Cheng, F., & Peng, D. (2020). Seismic Fault Interpretation Using Deep Learning-Based Semantic Segmentation Method. IEEE Geoscience and Remote Sensing Letters.

[17]. Yamagishi, K., Hori, J., & Miyakawa, M. (2006, August). Development of EOG-based communication system controlled by eightdirectional eye movements. In 2006 International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 2574-2577). IEEE.

[18]. Betke, M., Gips, J., & Fleming, P. (2002). The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. IEEE Transactions on neural systems and Rehabilitation Engineering, 10(1), 1-10.

[19]. Jacob, R. J. (1991). The use of eye movements in human-computer interaction techniques: what you look at is what you get. ACM Transactions on Information Systems (TOIS), 9(2), 152-169.

[20]. Królak, A., & Strumiłło, P. (2012). Eye-blink detection system for human–computer interaction. Universal Access in the Information Society, 11, 409-419.

[21]. Khare, V., Krishna, S. G., & Sanisetty, S. K. (2019, March). Cursor control using eye ball movement. In 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (Vol. 1, pp. 232-235). IEEE.

[22]. Toyama, K. (1998). Look, ma-no hands! hands-free cursor control with real-time 3d face tracking. PUI98.

[23]. Pradhan, K., Sayyed, S., Karhade, A., Dhumal, A., & Shaikh, S. EYEBALL MOVEMENT BASED CURSOR USING MACHINE LEARNING.