The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
19(2): S2: 287-294, 2024
www.thebioscan.com

Nose-Driven Cursor Control: An Assistive System for Disabled Individuals Using OCRM Similarity Tracking

[1]Gneneswara Rao Nitta, [2]Sai Likhitha Alapati , [3]Sneha Nitta,[4]Hari Prasada Raju

[1]Professor of CSE Department, DRK Institute of Science &Technology, Hyderabad, India,

[2]CSE Department, DRK Institute of Science &Technology, Hyderabad, India.

[3]Systems Operations Engineer, Wells Fargo India Solutions Pvt Ltd, Madhura Nagar Colony, Gachibowli, Hyderabad-500032, Telangana, India.

[3]Executive Director, Wells Fargo India Solutions Pvt Ltd, Madhura Nagar Colony, Gachibowli, Hyderabad-500032, Telangana, India.

[1]gnani.nitta@gmail.com, [2]alapati.sl2003@gmail.com, [3]siri.nitta24@gmail.com, [4]khariprasadaraju@yahoo.co.in

DOI: https://doi.org/10.63001/tbs.2024.v19.i02.S2.pp287-294

KEYWORDS
Similarity Measure,
OCRM,
Mouse operations,
Vision Computer,
Visual Tracking approach,
Communication devices.
Received on:

20-07-2024

Accepted on:

04-11-2024

ABSTRACT

This paper introduces a visual mouse system designed to assist disabled individuals in controlling computer mouse functions, such as cursor movement and clicking actions, using simple body gestures. Specifically, the system tracks the user's nose tip to execute mouse commands, providing an untouchable alternative for desktop interaction. The main goal is to create an accessible, low-cost solution using widely available hardware to make computing more inclusive for individuals with physical disabilities. Using a camera, the system captures and translates user head movements into mouse gestures on the screen. The proposed methodology is built around visual tracking, isolating a sample region from each video frame and then identifying that region in subsequent frames through Optimal Cost Region Matching (OCRM). This similarity measure ensures accurate and seamless tracking of the user's movement, providing an effective control mechanism. Experimental results demonstrate that the OCRM-based visual mouse system outperforms existing Camera Mouse applications regarding reliability and efficiency. The proposed approach delivers a flexible, cost-effective, and consistent solution for enhancing computer accessibility for disabled users.

INTRODUCTION

The world population is about 500 million and above, among which 10% of people are suffering from several certain kinds of disabilities. From the existing surveys, it was observed that diverse, supportive systems and devices are developed to assist the incapable and disabled individuals having extreme complexities in communicating their desires, feelings, and necessities. They use restricted volunteer movements to interact with family, associates, and caretakers. For Instance, specific individuals could move their heads, certain could flicker or wink voluntarily, and certainly could move their eyes or tongue, etc,. One of the foremost assisting tools for these disabled individuals is gadgets for communication like computers or similar devices. Nevertheless, the foremost issue with these devices is to monitor the complexities that are restricted merely to the physical talents of the individual. Thus, several interfaces are being constructed to enable interactions between disabled individuals and devices like computers [2-6].

The interfaces are classified into two groups comprising intrusive and non-intrusive approaches. Intrusive approaches frequently employ connecting instruments that measure personal actions and reactions. Even though intrusive approaches could identify features or indications more precisely, they require costly instruments that are not flexible [7-11]. Non-intrusive approaches frequently track individual gesticulations by processing images or videocassettes through the camera. Apart from the intrusive approaches, they are flexible for individuals and include a few expensive communication devices. The Supportive technological systems are presented to support disabled individuals and employ their volunteer motions to monitor the computers.

Individuals with disabilities benefit from computer access in numerous alternative manners, along with the improved ability to interact. These make the individuals more energetic, participating in entertaining events, using the internet, and using computer monitoring approaches like computerized wheelchairs. For instance, a mouse is one of the accessible devices that assist disabled individuals. These devices use ultraviolet emitters attached to an individual's eyeglasses, head cover, or a lid. These devices put the receiver above the screen and exploit an ultraviolet reflector connected to an individual's forehead or eyeglasses [12]. The people's head motions regulate the operations of the mouse pointer on the monitor. Mouse clicks are produced with the actual button or a software interface.

As an alternative to the existing accessible assisting devices, we aim to introduce a non-intrusive, flexible, consistent, and economical interacting instrument that is effortlessly adjustable to assist the necessities of disabled users. While matching with other monitoring tools that assist individuals with austere in-capabilities, these devices are luxurious, i.e., no physical body connections, flexible i.e., no adjustment, and simple that tracks numerous body parts. This tracking capability enhances functionalities for individuals who cannot make volunteer head motions.

The suggested methodology with mouse operation is a non-intrusive approach that assists disabled individuals in communicating with desktops and PCs. The nose is an appropriate tracking part that acts as a mouse pointer. Initially, it is flexible for any user to put their nose at a specific location while observing the monitor. The nose is precisely in the middle of the face and is not blocked whenever the individual head moves considerably. Subsequently, the nose pattern inclines to include a reasonable degree of illuminating divergence to its neighboring parts. In this environment, the nose inclination is brighter than the remaining part since it is tilted and leaned toward the overhead lighting.

The camera mouse device typically comprises more than one camera for seizing video frames and a processing element such as a desktop that employs the image processing approaches to transform a moving action into video frames to function as a mouse. The strategy is typically developed using visual tracking and a mouse-controlling unit. This tracking unit regains the moving data from the video, and the mouse regulating unit defines specific guidelines for monitoring. The Optimal Cost Region Matching [1] Similarity measure measures the similarities between the sub-image in the explorer window and the sample image picked from the preceding frame in the visual tracking unit.

A brief introduction to assistive technology and its use for severely disabled and disabled people is specified in this section. Section 2 deliberates the previous Assistive technology and its usage for Computer Mouse. The proposed OCRM-based tracking algorithm for mouse operation methodology is briefly defined in Section 3. Section 4 provides the Experimental outcomes. Section 5 concludes the suggested approach performance for better Assistive technology service.

1. Literature Survey

Dissimilar Approaches are suggested for the execution of camera mouse where maximum individuals employ head position or motions and face parts. Devices employing eye movements are previously given in [13, 14]]; here, the mouse actions and click events are engendered depending on eye gestures through fitting at a particular location for more than one second. Other alternative approaches for camera mouse execution are employing the mouse movements and lip interpretation suggested in [15, 16]. In this approach, diverse mouth shapes are utilized to read the lip. The approach suggested in [17] can monitor the mouse depending on eye and mouth actions. In [18], an approach motivated through visual face tracking is presented depending on a 3D prototype. This approach is essential to initiate with facial events elements in 3D space and formerly evaluates the revolution and transformation matrix. Nostrils and nose tip methods are suggested in [19, 20], which are correspondingly additional devices for controlling the mouse cursor.

Recently, a straddling ultraviolet radiating mechanism, a “comparative” pointing tool, has been acting similarly to a joystick compared to the mouse [21]. In [22], an ultraviolet transmitter is introduced, attached to the individual glasses, a group of ultraviolet acknowledged units that replace the computer's keyboard's values, and a tongue touch pane to stimulate the ultraviolet ray. A non-interacted, ultraviolet-dependent device that traces replicated laser speckle form of skin is suggested in [23]. Current investigation progress promises less luxurious gaze-tracking results [24]. Individuals with austere incapacities are frequently not capable of retaining their head motions to employ profitable gaze trackers consistently. Other controlling tools measure the electrooculography potential (EOG) to identify eye motions [25, 26] or investigate characteristics in electroencephalograms [27]. In [26], “Eagle Eyes,” i.e., EOG-aided devices are developed that permit individuals who could move their eyes to regulate the mouse. The testaments in [26] demonstrate that Eagle Eyes has significantly improved youngsters’ lives. Till now, certain children do not prefer to attach the electrodes to their faces. The other drawback of EOG-aided devices is that conductors can drop when the individual exudes.

2. Proposed Methodology

The proposed assisting system tracks the body part, i.e., the tip of the nose, using a camera and identifies movement to monitor the computer directly as a mouse cursor. The architectural representation of the proposed approach is given in Fig. 1. The Mouse operating device presently comprises two computers connected as a “vision computer” and an “individual computer.” The vision computer implements a visual tracing approach and directs the location of the traced part to every individual computer. The separate computer interprets obtained indications and executes application software that a person desires to perform. These two computers' functions are incorporated into a single computer; however, the present architecture guarantees adequate processing strength for visual tracking and allows an individual to control its tracking efficiency, deprived of intruding on individual events.

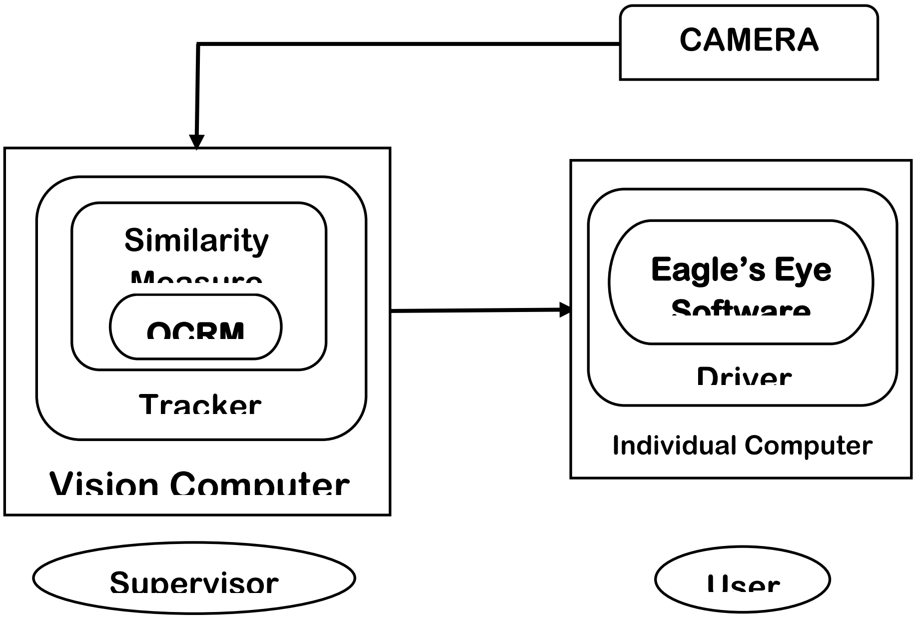


Fig. 1. The Camera Mouse system

2.1. Vision Computer

This system acquires and shows a video of an individual present ahead of the computer. This is accessed through a camera attached overhead or underneath the screen of the particular computer. Viewing this video, an individual or caretaker clicks using the mouse on the body part as an image to be traced. The camera's remote regulator is mainly used to alter the pan and slope positions of the camera and its zoom so that the preferred body part is aligned. The vision system defines the positions of designated part in the preliminary image and formerly evaluates automatically in succeeding images. The location of the traced part in every image frame is given to individual computers.

2.2. Individual Computer

The individual computer implements economical or customer software appliances a person selects. It executes a specific driving program in the background that employs the indication obtained from the vision computer, balances to synchronize the existing monitor resolution, and formerly replaces them with the locations of the pointer. The driving program depends on software built by the Eagle Eyes system [26] and executes autonomously from the individual preferred appliances. The Camera Mouse works as an actual mouse, and a physical button is given in Fig. 1, which is employed to work as a standard mouse. The individual moves this mouse cursor by moving the nose or any alternatively picked part. The driving program comprises alterations for horizontal and vertical "profit" aspects. A higher profit aspect makes the minor motion of the head move the mouse cursor at higher distances but with lesser accuracy.

There is a resemblance between altering profits and altering the zoom lens. The traced regions' positions are mounted through a specific amount whenever the profit changes. Varying the camera's zoom allows the vision computer to trace preferred regions with less or more information. If the camera lens is zoomed in on an area, this will incorporate a better percentage to control. Therefore, minor motions through the individual would exhibit higher motions of the pointer. On the other hand, if the camera lens is zoomed out, the region would incorporate a lesser percentage of the monitor, and therefore, bigger movements would be essential to move the pointer.

2.3. Tracking Approach

Whenever an individual primarily clicks on the region to be traced, it is observed that a sub-image inside the square is obtained by drawing a square around the region. The detached sub-image is called a "sample" to define the region's location in the subsequent image frame. To discover this location, the tracking approach employed this sample to explore the region in an exploration window, which is aligned at the center position of the region in the earlier frame. The sample is changed through this exploration window associated with principal sub-images. The window specified to comprise centers of entire sub-images is verified. Fig. 2 represents the sample and exploring window locations.

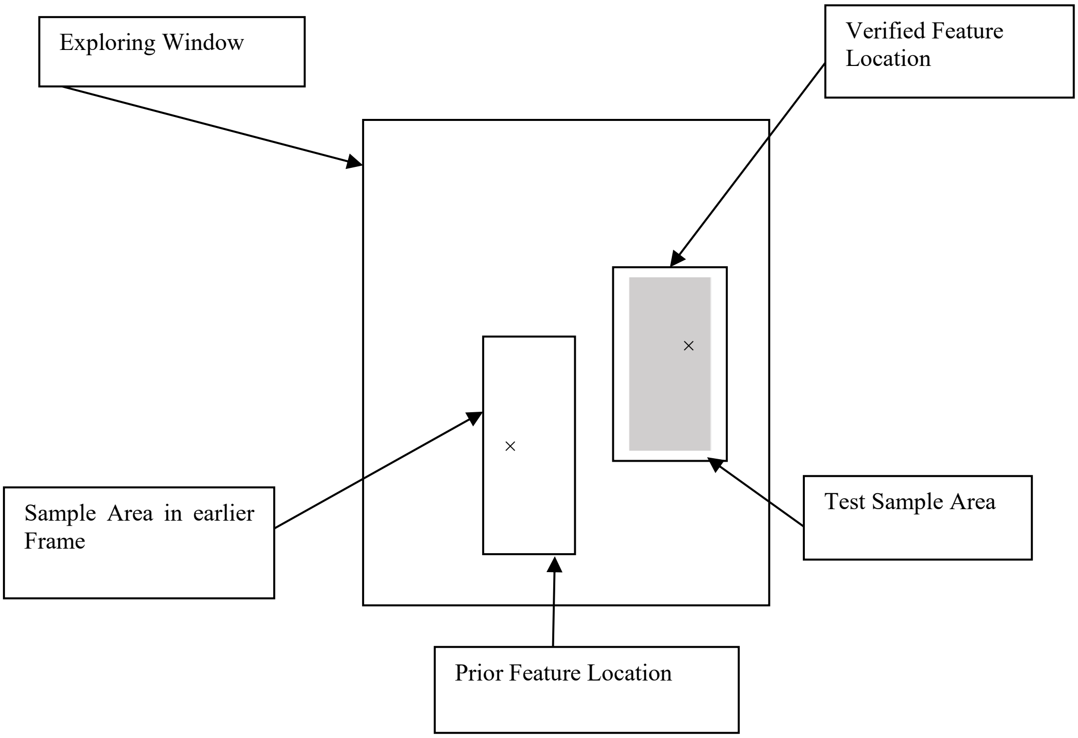


Fig2: Samples with the Exploration of the Search Window

Sequentially, for every 30th of a second, a new frame is obtained; the sample is typically identical to the illuminating feature structure in a novel frame that could be discovered by exploring for the finest associated sub-image. With the notion that the subsequent frames with consistent illuminating patterns are constant, the "constant illumination statement" is frequently given whenever developing an approach for image motion analysis. Every sub-image in the exploring window is compared with its sample sub-image, detached from the earlier frame. OCRM is employed for efficient image retrieval matching as a measure of comparison, which fulfills the personal insight in an exact manner. The sub-image picked from the existing frame has been aided as the sample to explore similar sub-image in subsequent frames. This procedure iterates, and the sample is upgraded for every frame. This sample update guarantees that robust similarity could be generated in every frame, along with tracking the movement of the primarily picked sample.

2.3.1. Optimal Cost Region Matching (OCRM) Similarity Measure

In the OCRM [1] approach, the sub-image and sample image are matched depending on the static number of blocks by attaining each block's feature vector. This size of each block is employed to augment the efficiency of color and textual features. It evaluates six features for every block, where three among them have the typical color elements in 4 × 4 block, and the remaining three specify energetic in higher occurring sets of wavelet transformations [28] such as the square root of 2nd order instantaneous of wavelet constants. The LUV color space [29] is exploited to mine color elements where L denotes the luminance, and U and V refer to the color data.

The subsequent phase segregates the image as blocks with 64 regions, each about mined color and textual features similar to the current image. In this technique, several regions are static to fulfill the characteristics of the Monge matrix. The segmentation of images depends on color and frequency, which are employed through the k-mean approach [30, 31]. The k-means approach separates the feature vectors into clusters related to a single area of the segmented image. The foremost benefit of employing the k-means clustering approach for division is that the blocks of every region are not adjacent.

The features from corresponding regions are considered for evaluation in the Region Descriptors. These regions are arranged in a non-descending order depending on the region descriptor. There is a restructuring in the regions, and the formal distance estimated amongst query and target images hinges on the given equation (2).

$$d_{i,j} = |q_i - t_i| \qquad (2)$$

here $d_{i,j}$ is the distance amongst the region descriptor of query and target image, q_i is region descriptor of ith region of the query image and t_i is a region descriptor of the jth region of the target image.

$$d_{i,j} + d_{i+1,j+1} \leq d_{i,j+1} + d_{i+1,j} \qquad (3)$$

Rendering to the Monge property, Equation 3 is constantly correct. Consequently, $d_{i,j}$ acquired after re-structuring the regions in the arranged way that continually fulfills the Monge property. Whenever the distance matrix fulfills this property, the North-West Corner rule is an approach that provides optimum results with a lesser evaluation period. This guideline gives an optimal result of transportation problems for entire vectors. Thus, this similarity measure consecutively obtains sub image identical to the detached sample image and tracks the sample for mouse movements.

3. Experimental Results And Analysis

Several body parts were verified to discover points that the individual could flexibly move and consistently track through the structure. As the presence of body parts varies amongst persons, tracking performance fluctuates amongst individuals. Numerous features, nevertheless, were consistently traced diagonally through the investigation group. Fig 3 demonstrates the interface for the suggested methodology where the nose tip was validated as a part of the body.

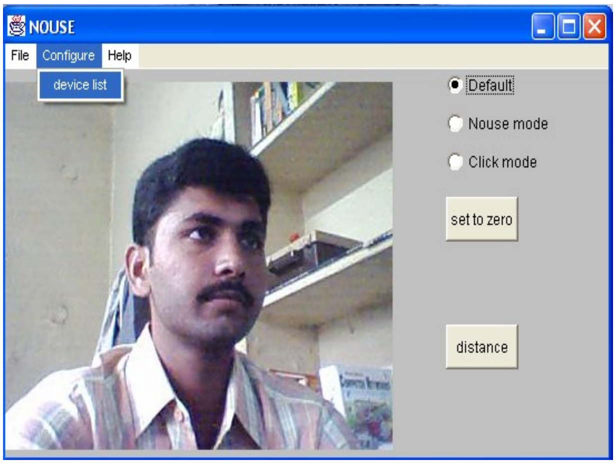
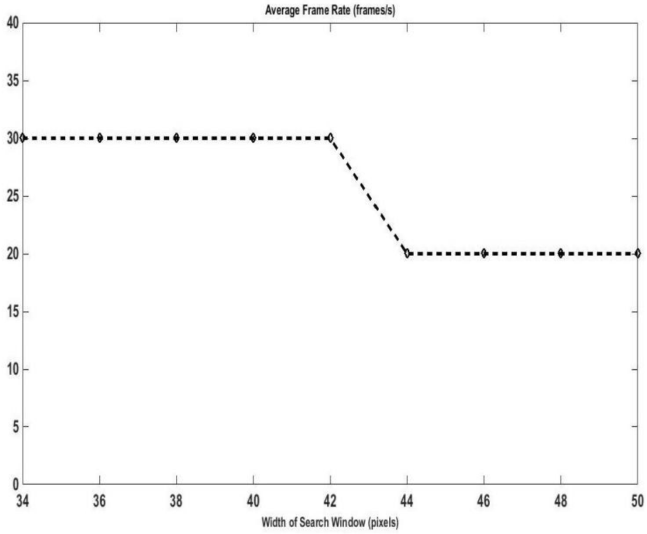


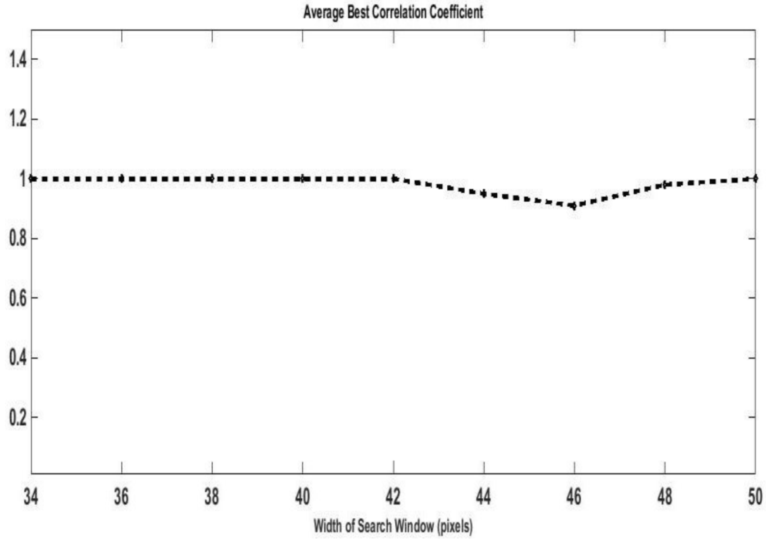
Fig 3: Interface for the proposed Methodology

Choice of parameters

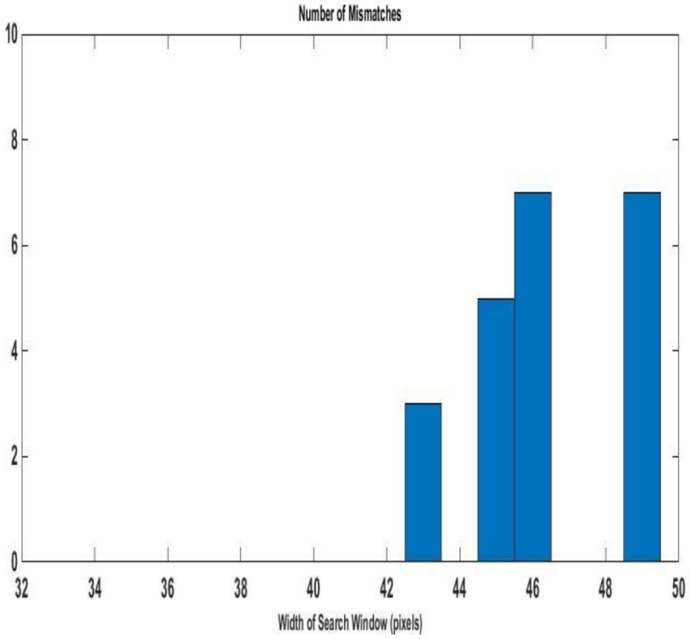
The sample dimension is 15 X 15 pixels, which is large enough to seize an eloquent body part and small enough such that a 40 X 40-pixel window could be explored at a frame rate of 30 Hz.



(a)

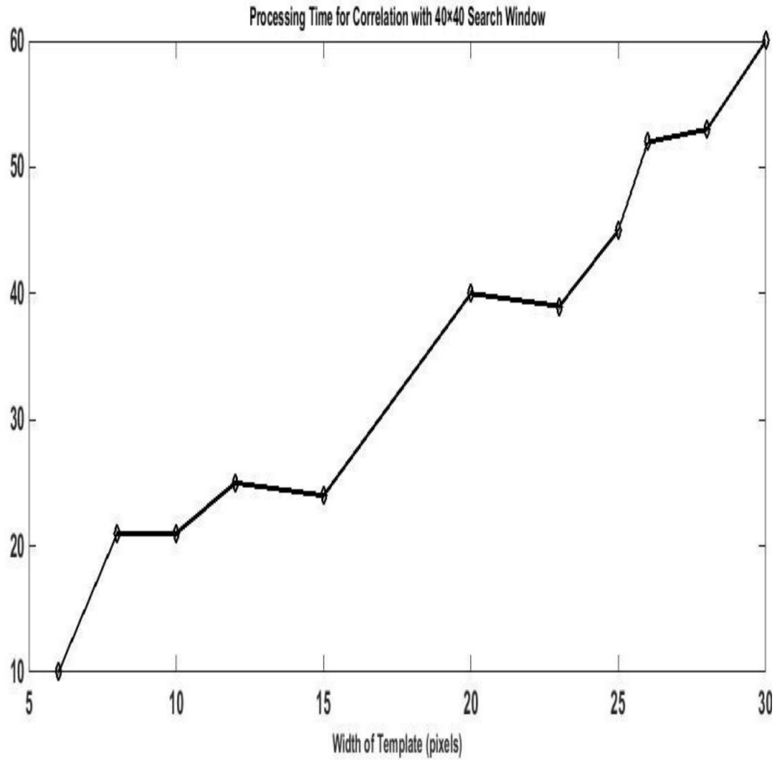


(b)

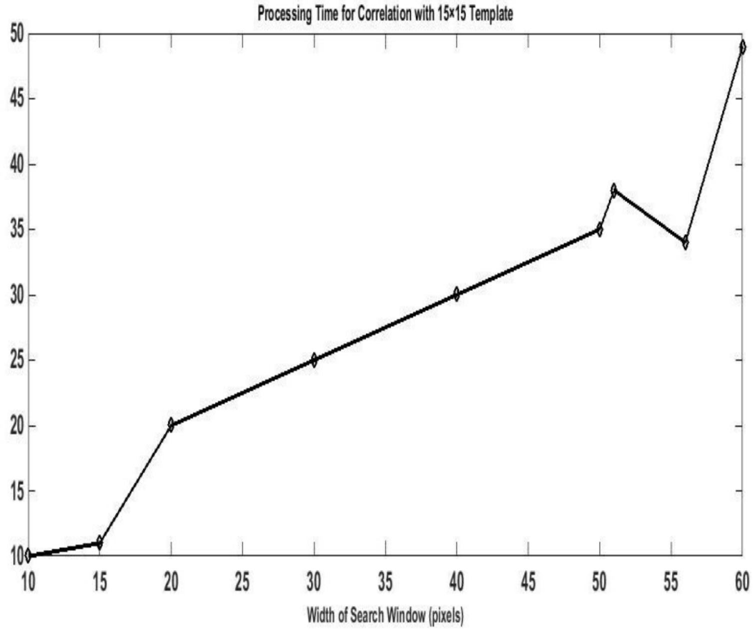


(c)

Fig 4. a) Rise in the width of the exploring window Vs the frame rate. b) Varying of the average correlation coefficient with varying window size. c) Graph between the width of the search window and the no of mismatches



(a)



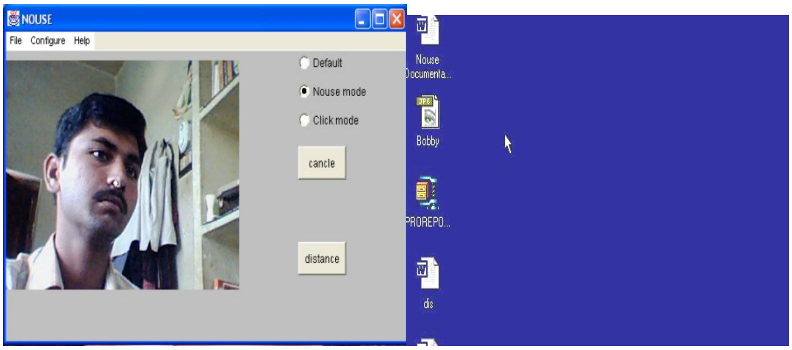
(b)

Fig 2.4. (a) The normalized correlation coefficient for a specified exploring window width of 40 pixels. (b) The normalized correlation coefficient for a specified 15X15 size sample

Figure 4(a) represents how a rise in the width of the exploring window influences the frame rate. Whenever the dimension of the window rises, the exploring space indirectly increases. Formerly, the time consumed to explore the window also rises. Thus, the number of frames processed within a second is spontaneously reduced. Figure 4(b) represents the varying average correlation coefficient when we change the search window size. When the size of the area increases, the area within which we are searching for the feature also increases. Then, the correlation coefficient value increases.

Figure 4(c) describes the graph between the width of the search window and the no of mismatches. When the size of the search window increases, no mismatches decrease. Up to a certain level, the no of mismatches decreases. We increased the search window size more than the processing time increased. The no frames processed within one second also decreases. Then, the feature goes out of the search window. So that no mismatches increase. Figure 5(a) shows the function period of the normalized correlation coefficient as a function of the width of the sample, specified an exploring window width of 40X40 pixels.

Figure 5(b) shows the functioning period of the normalized correlation coefficient as a function of the exploring window, specified as a fixed 15 X 15 size sample. The frame rate of 30 Hz = 33ms could be guaranteed if the individual dimensions of the selected sample and exploring window are 15 X 15 and 40 X 40 pixels. Fig 6 illustrates the selected feature's motions and the mouse's corresponding moment on the screen.



○ Default
● Nouse mode
○ Click mode

Fig 6: Outputs showing the selected feature's moment and the mouse's corresponding moment on the screen.

CONCLUSION

This study presents an innovative, assistive visual mouse system designed to improve computer accessibility for disabled individuals using a non-contact control method. Specifically, the system leverages nose tip tracking as an intuitive method to move the cursor, using widely available hardware such as a webcam. The Optimal Cost Region Matching (OCRM) similarity matching algorithm ensures accurate and reliable tracking of user movements, transforming these into mouse gestures for effective screen interaction.

Experimental results demonstrate that the proposed OCRM-based solution offers enhanced accuracy and flexibility compared to existing Camera Mouse applications. It reduces mismatches and improves responsiveness, contributing to a seamless and practical hands-free interaction experience. By utilizing low-cost components and avoiding invasive attachments, this system provides an affordable and adaptable alternative to existing assistive technologies, enhancing the quality of life and independence of users with physical disabilities.

Future work could focus on further optimizing the tracking algorithm for varied lighting conditions, expanding the set of body features that can be tracked, and integrating advanced machine learning techniques to increase robustness and adaptability to different users' needs. Additionally, broader usability studies can be conducted to further validate the system's effectiveness across diverse populations of individuals with disabilities.

REFERENCES

- N. Gnaneswara Rao, T. Sravani and V. Vijaya Kumar, “OCRM: Optimal Cost Region Matching Similarity Measure for Region Based Image Retrieval “, International Journal of Multimedia and Ubiquitous Engineering, Vol.9, No.4, pp.327-342, 2014.
- D. Stefanov, Z. Bien, and W. Bang, “The Smart House for Older Persons and Persons with Physical Disabilities: Structure, technology arrangements, and Perspectives,” IEEE Transaction on Neural Systems and Rehabilitation Engineering, vol. 12, no. 2, pp. 228-250, 2004.
- Md Samiul Haque Sunny, Javier Sanjuan De Caro, Ivan Rulik, Md Ishrak Islam Zarif, Mohammad Rahman, Inga Wang, Sheikh Ahamed, Nose Tracking Assistive Robot Control for the People with Motor Dysfunctions, Archives of Physical Medicine and Rehabilitation, Volume 102, Issue 10, 2021, Pages e82-e83,
- https://doi.org/10.1016/j.apmr.2021.07.718.
- J. Alon, V. Athitsos, and S. Sclaroff, “Simultaneous localization and recognition of dynamic hand gestures,” in Proc. 2005 IEEE MotionWorkshop, 2005.
- C. S. Lin, C. W. Ho, C. N. Chan, C. R. Chau, Y. C. Wu, and M. S. Yeh, “An eye-tracking and head-control system using movement increment coordinate method,” in Proc. 2007 Optics & Laser Technology Conf., pp. 1218–1225.
- S. S. Khan, M. S. H. Sunny, M. S. Hossain, E. Hossain and M. Ahmad, "Nose tracking cursor control for the people with disabilities: An improved HCI," *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, Bangladesh, 2017, pp. 1-5, doi: 10.1109/EICT.2017.8275178.
- C. Mauri, T. Granollers, J. Lorés, M. García, “Computer vision interaction for people with severe movement restrictions,” An Interdisciplinary Journal on Humans in ICT Environments, vol. 2, no. 1, pp. 38-54, April 2006.
- G. M. Eom, K. S. Kim, C. S. Kim, J. Lee, S. C. Chung, B. Lee, H. Higa, N. Furuse, R. Futami, and T. Watanabe, “Gyro Mouse for the Disabled:‘Click’ and ‘Position’ Control of the Mouse Cursor,” InternationalJournal of Control, Automation, and Systems, vol. 5, no. 2, pp. 147-154, April 2007.
- C. Veigl, “An Open-Source System for Biosignal- and Camera-MouseApplications,” studying Medical Computer Science, TechnicalUniversity Vienna, 2006.
- C. Topal, A. Doğan, and Ö. N. Gerek, “A Wearable Head-MountedSensor-Based Apparatus for Eye Tracking Applications,” in Proc.VECIMS 2008 IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems.
- M. A. Qamar, and A. Jehanzeb, “Retina Based Mouse Control(RBMC),” World Academy of Science, Engineering and Technology,2007.
- O. Takami, N. Irie, C. Kang, T. Ishimatsu, and T. Ochiai, “Computer interface to use head movement for handicapped people,” in Proc. IEEETENCON’96, Digital Signal Processing Applications, vol. 1, 1996, pp.468–472.
- K. E. Yi, and K. S. Kuk, “Eye tracking using neural network and mean-shift, ”LNCS, vol. 3982, pp. 1200–1209, 2006.
- J. J. Magee, M. R. Scott, B. N. Waber, and M. Betke, “EyeKeys: A real-time vision interface based on gaze detection from a Low-grade VideoCamera,” Computer Science Department, Boston University, 2004.
- H. E. Cetingul, Y. Yemez, E. Erzin, and A. M. Tekalp, “Discriminative analysis of lip motion features for speaker identification and speechreading,” IEEE Trans Image Process, vol. 15, no. 10, pp. 2879-91, 2006.
- S. Stillittano, and A. Caplier, “Inner lip segmentation by combining active contours and parametric models,” in Proc. of international conference on Computer Vision Theory and Applications., pp. 297-304.2008.
- Y. Shin, J. S. Ju, E. Y. Kim, “Welfare interface implementation using multiple facial features tracking for the disabled people,” PatternRecognition Letters, vol. 29, pp. 1784–1796, 2008.
- J. Tu, H. Tao, T. Huang, “Face as mouse through visual face tracking, ”Computer Vision and Image Understanding, vol. 108, pp. 35-40, 2007.
- C. M. Yee, J. Varona, and F. J. Perales, “Face-Based PerceptualInterface for Computer-Human interaction,” Department de Matemàtiques I Informàtica, 2006.
- J. Na, W. Choi, D. Lee, “Design and Implementation of a MultimodalInput Device Using a Web Camera,” ETRI Journal, vol. 30, no. 4, pp.621-623, August 2008.

- D. G. Evans, R. Drew, and P. Blenkhorn, “Controlling mouse pointer position using an infrared head-operated joystick,” IEEE Trans. Rehab.Eng., vol. 8, no. 1, pp. 107–117, 2000.
- Y. L. Chen, F. T. Tang, W. H. Chang, M. K. Wong, Y. Y. Shih, and T. S. Kuo, “The new design of an infrared-controlled human-computer interface for the disabled,” IEEE Trans. Rehab. Eng., vol. 7, pp. 474–481, Dec. 1999.
- R. B. Reilly and M. J. O’Malley, “Adaptive noncontact gesture-based system for augmentative communication,” IEEE Trans. Rehab. Eng., vol. 7, no. 2, pp. 174–182, 1999.
- C. H. Morimoto, D. Koons, A. Amit, M. Flickner, and S. Zhai, “Keeping an eye for HCI,” in Proc. XII Brazilian Symp. Computer Graphics andImage Processing, 1999, pp. 171–176.
- J. R. LaCourse and F. C. Hludik Jr., “An eye movement communication control system for the disabled,” IEEE Trans. Biomed. Eng., vol. 37, no.12, pp. 1215–1220, 1990.
- P. DiMattia, F. X. Curran, and J. Gips, An Eye Control TeachingDevice for Students Without Language Expressive Capacity: EagleEyes.Lampeter, U.K.: Edwin Mellen, 2001.
- M. Pregenzer and G. Pfurtscheller, “Frequency component selection for an EEG-based brain to computer interface,” IEEE Trans. Rehab. Eng., vol. 7, no. 4, pp. 413–419, 1999.
- I. Daubechies, “Ten Lectures on Wavelets”, Capital City Press, (1992).
- B. G. Prasad, S. K. Gupta and K. K. Biswas, “Region-Based Image Retrieval Using Integrated Color Shape and Location Index”, Computer Vision and Image Processing, vol. 94, (2004), pp. 193–233.
- J. Li, J. Z. Wang and G. Wiederhold, "Classification of textured and non-textured images using region segmentation", Proceedings of the Seventh International Conference on Image Processing, Vancouver, BC, Canada, (2000) September.
- B. S. Manjunath, et al., “Color and texture descriptors”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, (2001), pp. 703–715.