Controlling Mouse Navigation Through 3D Head Movement

Pradeep V, Jogesh Motwani

Abstract: Various results have been proposed in the past decades to capture the user's head motions through a camera to control the navigation of the mouse pointer to enable the people with disability in the movement to interact with computers. Movement of the facial feature is tracked to estimate the movement of the mouse cursor in the computer screen. Synchronizing the rate of movement of the head with the mouse cursor movement is identified as the challenge as the head movement is three dimensional but the sequence of images captured by the web camera is two dimensional. The proposed system is an innovative approach of capturing the three-dimensional head rotation through the usual web camera that captures the image in two dimensions.

Index Terms: 3D head movement, Assistive technology, Camera mouse, Controlling mouse cursor, Hands-free computing, People with disability in movement.

1 INTRODUCTION

ABOUT 5.4 million people in India have a disability in movement as per the census 2011 [1]. They have not received an equitable chance like others to access the computers easily due to their inconvenience of using the standard input devices of the computer. As on-screen virtual keyboards can be used to simulate the physical keyboard, an alternative solution to emulate the mouse cursor movement and click operations is highly desirable to support the people with disability in movement. Various results have been proposed in the past decades to capture the users' head motions through a camera to control the navigation of the mouse pointer to enable the people with disability in the movement to interact with computers. Palleja et al. [2], Kim et al. [3], Frank et al. [4], Epstein et al. [5] and John et al. [6] have proposed the mouse replacement solutions that translate the user's head movements to mouse cursor movements. Nabati et al. [7], Zhu et al. [8], Kumar et al. [9], Woramon et al. [10], Manresa [11] and Gyawal et al. [12] have presented an approach to control the mouse pointer by tracking the face region. Javier et al. [13], Chathuranga et al. [14], Bian et al. [15], Gorodnichy et al. [16], Mohamed et al. [17] and Morris et al. [18] have tracked nose region to control the mouse pointer. Chairat et al. [19] and Parmar et al. [20] have tracked the region between the eyes whereas Eric et al. [21] locates the tracking point near the upper lip to simulate the mouse cursor movement. Betke et al. [22], Akram et al. [23] and Connor et al. [24] use the feature selected by the attending care such as the tip of the nose, eye, lip and converts them into the mouse pointer movement on the screen. Yunhee et al. [25], Fathi et al. [26] and Kim et al. [27] have proposed the systems to implement the mouse cursor movement by tracking the users' eye movement whereas Magee et al. [28], Sugano et al. [29], Uma et al. [30], Valenti et al. [31] and M. Nasor et al. [32] have attempted tracking the eye gazes.

- *Pradeep V, Research Scholar, Department of CSE, Channabasaveshwara Institute of Technology, Tumkur, India. Visvesvaraya Technological University, Belagavi, India, PH-7406057060, Email: writetopv@gmail.com*
- *Jogesh Motwani, Professor, Department of CSE, Channabasaveshwara Institute of Technology, Tumkur, India. Visvesvaraya Technological University, Belagavi, India, PH-9487626085, Email: jogeshmotwani@gmail.com*

2 PROBLEM STATEMENT

Many proposed solutions have used OpenCV Haar Cascade object detection algorithm proposed by Viola and Jones [33] to capture the users' head motions through a camera for controlling the navigation of the mouse pointer in the computer monitor screen. The horizontal and vertical movement of the mouse cursor is controlled by the respective horizontal and vertical movement of the head. The rate of movement of the head will not synchronize with mouse cursor movement as the head movement is three dimensional but the sequence of images captured by the web camera is two dimensional. Fig. 1 shows the average variation in vertical position of the mouse cursor between the current frame and previous frame when the 2D head is captured and the head is moved from top to bottom in almost a constant rate. The users' head is captured as a rectangular subset of the image using Haar Cascade classifier algorithms of OpenCV. The mid-point of the rectangular subset is considered for calculating the variation.

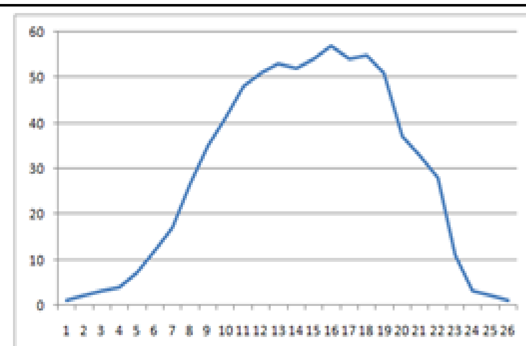

Fig. 1. *Average variation in the vertical position of the mouse cursor while capturing the head in 2 dimensional.*

Similarly, Fig. 2 shows the average variation in the horizontal position of the mouse cursor between the current frame and previous frame when the 2D head is captured and the head is moved from left to right in almost a constant rate.

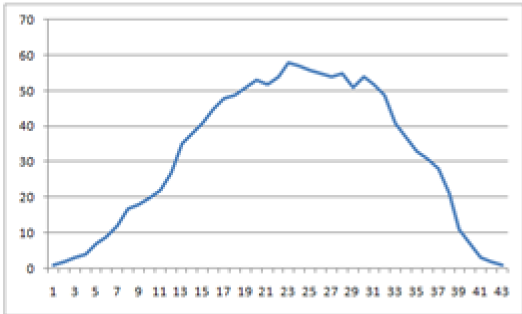

***Fig. 2**. Average variation in horizontal position of the mouse cursor while capturing the head in 2 dimensional.*

3 METHODOLOGY

To rectify this huge variation on synchronizing the rate of mouse cursor movement with the head movement, the proposed system captures the three-dimensional movement of the users' head through the two-dimensional sequence of images captured by the web camera. OpenCV Haar Cascade object detection algorithm proposed by Viola and Jones [33] is used to detect the face region. The ROI of the proposed system is estimated using the DLib library's built-in pre-trained facial landmark detector proposed by Kazemi and Sullivan [34] which recognizes 68 specific landmark points on a face. The following four ROIs of the proposed system on the face region are shown in Fig.3: 1. Inner left eye corner, 2. Inner right eye corner, 3. Nose bridge and 4. Nose tip. The four regions of the facial part are extracted from the facial landmark using the mapped indices 39, 42, 27 and 30 and the respective (x,y) coordinates are grabbed.

***Fig. 3**. ROIs on the face region.*

Capturing the vertical movement of the head is shown in Fig. 4. The nose bridge and the nose tip are the ROIs for capturing the vertical movement of the head. Increase in the vertical distance between the two ROIs indicates the head moving towards the bottom and the same will be reflected in the mouse cursor movement, moving down in the vertical direction. Similarly, a decrease in the above vertical distance will be used to move the mouse cursor upward in the vertical direction. Capturing the horizontal movement of the head in three dimensional is shown in Fig. 5. The inner left eye corner (X_1), nose tip (X_2) and inner right eye corner (X_3) are the ROIs for capturing the horizontal movement of the head. Increase in the horizontal distance between X_2 and X_3 and decrease in the horizontal distance between X_1 and X_2 indicates the head moving towards left and the same will be reflected in the mouse cursor movement, moving left in the horizontal direction. Similarly, decrease in the horizontal distance between X_2 and X_3 and increase in the horizontal distance between X_1 and X_2 indicates the head moving towards right and the same will be reflected in the mouse cursor movement, moving right in the horizontal direction.

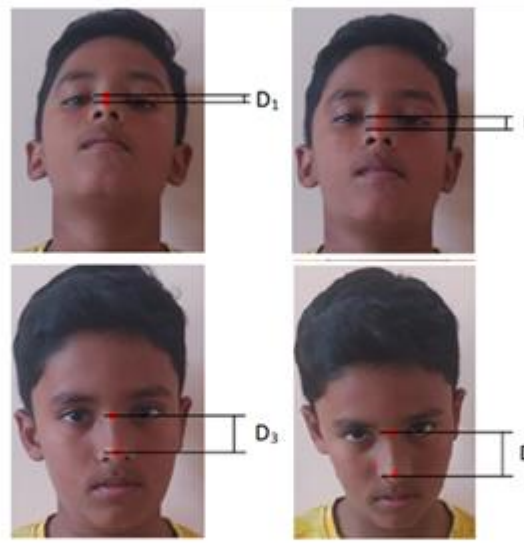

***Fig. 4**. Capturing the vertical movement of the 3D head.*

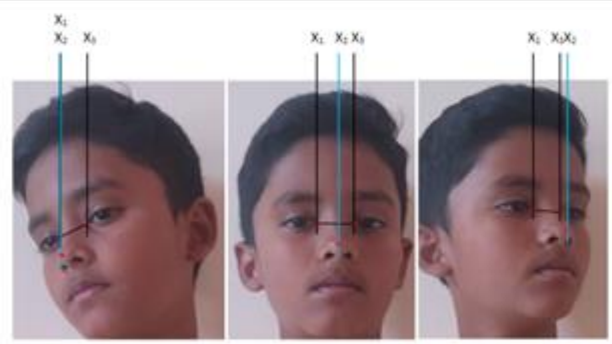

***Fig. 5**. Capturing the horizontal movement of the 3D head.*

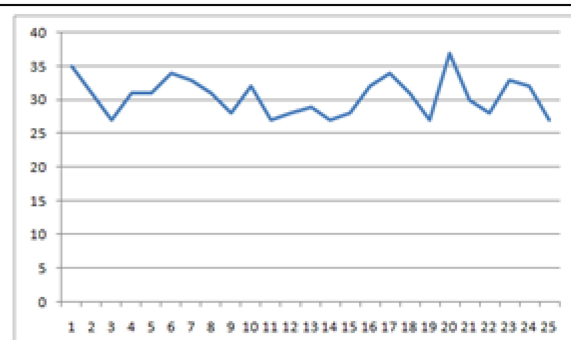

***Fig. 6**. Average variation inthe vertical directionwhile capturing the head in 3 dimensional.*

4 RESULT

The system is tested in the laptop computer with the following configuration:

- Intel Pentium CPU B950 2.10 GHz
- 2 GB RAM
- Logitech C170 Webcam
- Windows 7 Professional 32-bit Operating System
- 1366 x 768 Screen Resolution with Landscape orientation
- Microsoft Visual C++ 2008 Express Edition
- OpenCV 2.1.0
- Dlib C++ Library

Fig. 6 shows the average variation in the vertical direction of the 3D head captured between the current frame and previous frame when the head is moved from top to bottom in almost a constant rate.

Similarly, Fig. 7 shows the average variation in the horizontal direction of the 3D head captured between the current frame and previous frame when the head is moved from left to right in almost a constant rate.

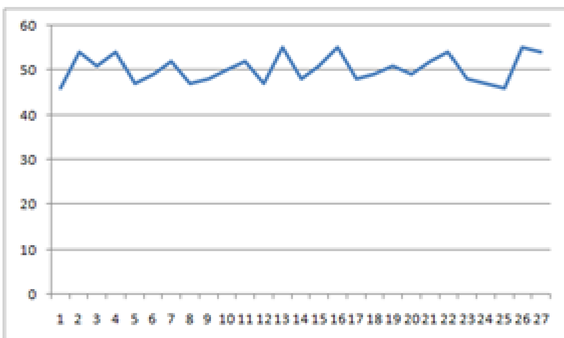

***Fig. 7**. Average variation in the horizontal direction while capturing the head in 3 dimensional.*

The variation in pixels of the successive head captured should be almost the same when the user moves the head in almost constant rate for successful synchronization. With the tolerance of 5 pixels of variation, i.e., the difference between the minimum and maximum threshold pixel value, the success rate of synchronization achieved is 64% (Fig. 8) in the vertical direction and 56% (Fig. 9) in the horizontal direction, while capturing the head in 3 dimensional, comparing to the success rate of synchronization achieved as 26% (Fig. 10) in the vertical direction and 25% (Fig. 11) in the horizontal direction, while capturing the head in 2 dimensional.

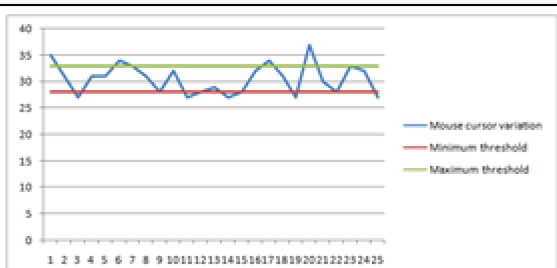

***Fig. 8**. 64% of synchronization in the vertical direction with the tolerance of 5 pixels of variation while capturing the head in 3 dimensional*

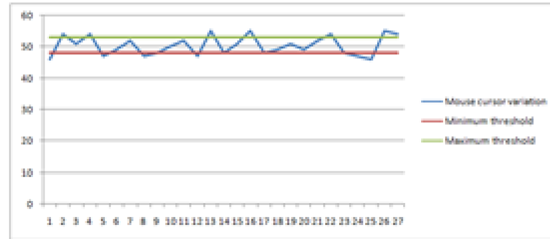

***Fig. 9**. 56% of synchronization inthe horizontaldirection with the tolerance of 5 pixels of variation while capturing the head in 3 dimensional.*

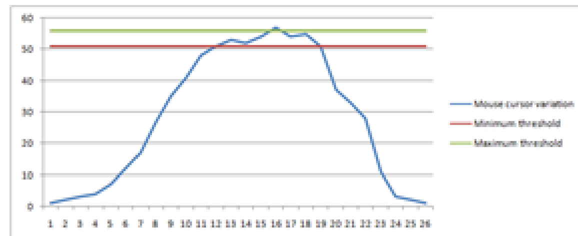

***Fig. 10.** 26% of synchronization inthe vertical direction with the tolerance of 5 pixels of variation while capturing the head in 2 dimensional.*

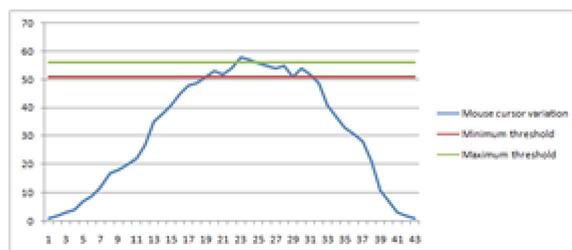

***Fig. 11**. 25% of synchronization in the horizontal direction with the tolerance of 5 pixels of variationwhile capturing the head in 2 dimensional.*

With the tolerance of 10 pixels of variation, the success rate of synchronization achieved is 96% (Fig. 12) in the vertical direction and 100% (Fig. 13) in the horizontal direction, while capturing the head in 3 dimensional, comparing to the success rate of synchronization achieved as 33% (Fig. 14) in the vertical direction and 35% (Fig. 15) in the horizontal direction, while capturing the head in 2 dimensional.

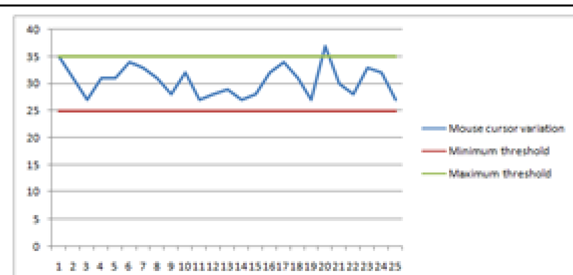

***Fig. 12**. 96% of synchronization in vertical direction with the tolerance of 10 pixels of variation while capturing the head in 3 dimensional.*

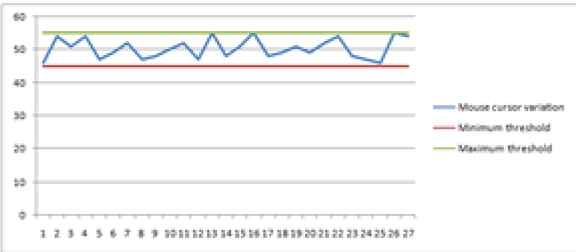

***Fig. 13**. 100% of synchronization inhorizontal direction with the tolerance of 10 pixels of variation while capturing the head in 3 dimensional.*

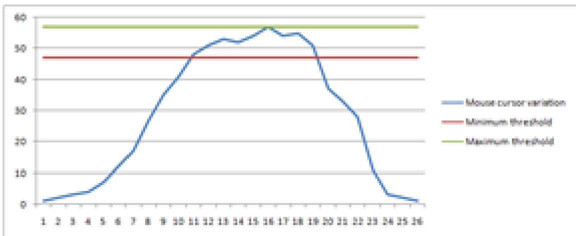

***Fig. 14**. 33% of synchronization in vertical direction with the tolerance of 10 pixels of variation while capturing the head in 2 dimensional*

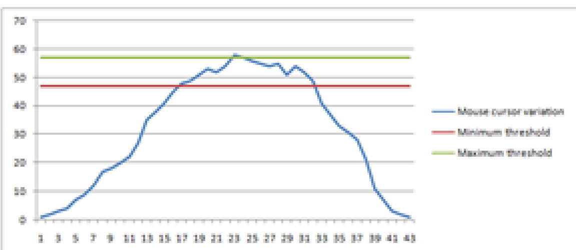

***Fig. 15**. 35% of synchronization in horizontal direction with the tolerance of 10 pixels of variation while capturing the head in 2 dimensional.*

4 CONCLUSION

This paper is focused on capturing the three-dimensional head rotation through the usual web camera that captures the image in two dimensions to achieve the synchronization of the mouse cursor movement with the users' head movement. The same or better performance can be achieved by placing luminous stickers of different colors on the four ROIs proposed on the users' face region. But, the main intention of the proposed system is to avoid such highly intrusive requirements for the people with disability in movement.

REFERENCES

[1] Social Statistics Division, Ministry of Statistics and Programme Implementation, Government of India, New Delhi (2017) Disabled Persons in India, A statistical profile 2016, http://www.mospi.gov.in.

[2] T. Palleja, E. Rubion, M. Teixido, M. Tresanchez, A. Fernandez del Viso, C. Rebate, J. Palacion, "Simple and Robust Implementation of a Relative Virtual Mouse Controlled by Head Movements" IEEE Conference on Human System Interaction, pp 221-224, 25-27, May 2008.

[3] Kim H, Ryu D (2006) Computer control by tracking head movements for the disabled. In: 10th International conference on computers helping people with special needs (ICCHP), Linz, Austria, LNCS 4061. Springer, Berlin, pp 709–715.

[4] Frank Loewenich , Frederic Maire, Hands-free mouse-pointer manipulation using motion-tracking and speech recognition, Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces, November 28-30, 2007, Adelaide, Australia.

[5] S. Epstein, E. Missimer and M. Betke. "Using kernels for a video-based mouse-replacement interface," Personal and Ubiquitous Computing, 18(1):47-60, January 2014.

[6] Magee, John; Epstein, Samuel; Missimer, Eric; Kwan, Christopher; Betke, Margrit. "Adaptive mouse-replacement interface control functions for users with disabilities", Technical Report BUCS-TR-2011-008, Computer Science Department, Boston University, March 2, 2011.

[7] Nabati, M., Behrad, A.: 3D Head pose estimation and camera mouse implementation using a monocular video camera. Signal Image Video Process. (2012, December).

[8] Zhu Hao, Qianwei Lei, J. Clerk Maxwell, "Vision-Based Interface: Using Face and Eye Blinking Tracking with Camera", Proceedings of the 2nd International Symposium on Intelligent Information Technology Application, vol. 2, no. 1, pp. 68-73, 2008.

[9] S. Kumar, A. Rai, A. Agarwal, N. Bachani, "Vision based human interaction system for disabled", 2Nd IEEE International Conference On Image Processing Theory Tools And Applications, 2010.

[10] Woramon Chareonsuk, Supisara Kanhaun, Kattaleeya Khawkam and Damras Wongsawang, "Face and Eyes mouse for ALS Patients", 2016 Fifth ICT International Student Project Conference (ICT-ISPC).

[11] C. Manresa-Yee, J. Varona, T. Ribot, F.J. Perales, "Non-verbal communication by means of head tracking", Ibero-American Symposium on Computer Graphics - SIACG (2006).

[12] P. Gyawal, A. Alsadoon, P. W. C. Prasad, L. S. Hoe, A. Elchouemi, "A novel robust camera mouse for disabled people (RCMDP)", 2016 7th International Conference on Information and Communication Systems (ICICS), pp. 217-220, 2016.

[13] Javier Varona, Cristina Manresa-Yee, Francisco J. Perales, "Hands-free vision-based interface for computer accessibility", Journal of Network and Computer Applications, Volume 31, Issue 4, November 2008, Pages 357-374.

[14] S. K. Chathuranga, K. C. Samarawickrama, H. M. L. Chandima, K. G. T.D. Chathuranga, A. M. H.S. Abeykoon, "Hands free interface for Human Computer Interaction", 2010 Fifth International Conference on Information and Automation for Sustainability, pp. 359-364, 2010.

[15] Z.-P., Bian, J. Hou, L.-P., Chau, N. Magnenat-Thalmann, "Facial position and expression based human computer interface for persons with tetraplegia", IEEE J. Biomed. Health Informat., vol. 20, no. 3, pp. 915-924, May 2016.

[16] Gorodnichy, D. and Roth, G., Nouse 'Use your nose as a mouse' perceptual vision technology for hands-free games and interfaces. In Image and Video Computing, Volume 22, Issue 12, Pages 931- 942, NRC 47140, 2004.

[17] A. Mohamed, R. Koggalage, "Control of mouse

movements using human facial expressions", Third IEEE International Conference On Information And Automation For Sustainability, pp. 13-18, 2007.
[18] T. Morris, V. Chauhan, Facial feature tracking for cursor control, Journal of Network and Computer Applications, v.29 n.1, p.62-80, January 2006.
[19] Chairat Kraichan and Suree Pumrin, “Face and Eye Tracking For Controlling Computer Functions”, 2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).
[20] K. Parmar et al., "Facial-feature based Human-Computer Interface for disabled people", Communication Information & Computing Technology (ICCICT) 2012 International Conference on, pp. 1-5, 2012.
[21] Eric Missimer , Margrit Betke, Blink and wink detection for mouse pointer control, Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, June 23-25, 2010, Samos, Greece.
[22] M. Betke, J. Gips, P. Fleming, "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, vol. 10, pp. 1-10, 2002.
[23] Akram W., Tiberii L., Betke M. (2007) A Customizable Camera-Based Human Computer Interaction System Allowing People with Disabilities Autonomous Hands-Free Navigation of Multiple Computing Tasks. In: Stephanidis C., Pieper M. (eds) Universal Access in Ambient Intelligence Environments. Lecture Notes in Computer Science, vol 4397. Springer, Berlin, Heidelberg.
[24] Connor C., Yu E., Magee J., Cansizoglu E., Epstein S., Betke M. (2009) Movement and Recovery Analysis of a Mouse-Replacement Interface for Users with Severe Disabilities. In: Stephanidis C. (eds) Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments. UAHCI 2009. Lecture Notes in Computer Science, vol 5615. Springer, Berlin, Heidelberg.
[25] Yunhee Shin, Eun Yi Kim, Welfare interface using multiple facial features tracking, Proceedings of the 19th Australian joint conference on Artificial Intelligence: advances in Artificial Intelligence, December 04-08, 2006, Hobart, Australia.
[26] Fathi, A. & Abdali-Mohammadi, “Camera-based eye blinks pattern detection for intelligent mouse”, Signal Image and Video Processing, November 2015, Volume 9, Issue 8, pp 1907–1916.
[27] Kim E.Y., Kang S.K. (2006) Eye Tracking Using Neural Network and Mean-Shift. In: Gavrilova M. et al. (eds) Computational Science and Its Applications - ICCSA 2006. ICCSA 2006. Lecture Notes in Computer Science, vol 3982. Springer, Berlin, Heidelberg.
[28] 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 Video Camera", In Proceedings of the IEEE Workshop on Real-Time Vision for Human-Computer Interaction (RTV4HCI), Washington, D.C., July 2004.
[29] Y. Sugano, Y. Matsushita, Y. Sato, H. Koike, "Appearance-based gaze estimation with online calibration from mouse operations", IEEE Transactions on Human-Machine Systems, vol. 45, no. 6, pp. 750-760, 2015.
[30] Uma Sambrekar, Dipali Ramdasi, "Estimation of Gaze For Human Computer Interaction", International Conference on Industrial Instrumentation and Control (ICIC), pp. 1236-1239, 2015.
[31] R. Valenti, N. Sebe, T. Gevers, "Visual gaze estimation by joint head and eye information", ICPR, 2010.
[32] Nasor, M., Rahman, K. K. M., Zubair, M. M., Ansari, H. and Mohamed, F. (2018) Eye-controlled mouse cursor for physically disabled individual, Proceedings of International Conference on Advances in Science and Engineering Technology, Abu Dhabi, UAE, pp. 152-156, 6 Feb-5 Apr, IEEE.
[33] Paul Viola, Michael Jones, (2001), Rapid Object Detection using a Boosted Cascade of Simple Feature, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, Vol. 1 pp. I-511 - I-518
[34] Kazemi, Vahid, and Josephine Sullivan. "One millisecond face alignment with an ensemble of regression trees." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.