iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** III **Month of publication:** March 2015

DOI:

www.ijraset.com Volume 3 Issue III, March 2015
IC Value: 13.98 ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Haar Based Head Movement Tracking with Voice commands for Computer control

Devanand G. Khandar[#1], Prof. R.V. Mante[#2], Prof. Dr. P.N. Chatur[#3]
[#] *Computer Science and Engineering GCOEA, SGBAU Amravati*

***Abstract*— This Paper, Mouse movement is done by vision based Head tracking. The centre of head detection is mapped to the mouse coordinates in computer screen. Left click and right click are controlled by voice left and right click command, respectively. Computer Vision has increased the size of research aiming to provide robust and effective techniques of real-time head movement detection and tracking. It provides alternative solutions for convenient device control, which confidences the application of interactive computer games, machine learning and guidance, robotic arm control, and machine access for disabled people who are physically challenged to do day to day life activity. For Head movement detection, several different approaches have been proposed and used to implement different algorithms for these technologies. Main objective is to overcome the limitations of the existing system by providing efficient touch less environment to the user, access computer's content via user's speech and different types of cursors, provide interactive visual sensing gaming.**
***Keywords*—Head Tracking, Human Computer Interaction, Computer Vision, Voice Recognition, Haar.**

I. INTRODUCTION

A Human life getting improves by the use of a recent new application in the field of Human-Computer Interaction (HCI) and Computer Vision (CV). To deal with the versatile use cases is requiring the future desired Perceptual User Interfaces (PUI) and user interface devices for more reliable and fast system performance. Nowadays much attention has been focused on the scenario of tracking user's movements with a video camera and translating the motion parameters into semantic indicatory symbols to manipulate machine operation. Head and its movements are important in expressing a person's desires, needs and emotional state. Head movement is also found to be a natural, simple and effective way of pointing to objects, interaction and communication. Thus, head movement detection has received significant attention in recent research. One of the various purposes for head movement detection and tracking is to allow the user to interact with a computer [1]. It also provides the ability to control many devices by mapping the position of the head into control signals.

Head movements are the least affected by disabilities because, for example, spinal cord injuries do not affect the ability to control them, as they are directly controlled by the brain [1] and movement is natural. Combining voice commanding and head movement detection can provide a larger number for possible control commands to be used with assistive technologies such as a wheelchair. Interaction with a computer is one of the various purposes allowed to the user for head movement detection and tracking. Vision based perceptual user system provides an alternative solution for convenient device control. It also significantly helps the disabled and elder people use limited voluntary actions to communicate with others [2].

A hand-free PUI that tracks human body movement in video to manipulate virtual computer pointing devices to respond user intentions is represented by a camera mouse system. Over the last decade, active researches have proposed to navigate cursor and trigger mouse click with the movement of eyes, nose, and face [2]. These systems require reliable and fast face tracking strategy, which allows users the comfortable and sufficient motion control. Head tracking can neither automatically initialize itself, nor handle large scaling case. Cursor navigation [3] is handled by face localization. We calculate the relative position of tracking window in the image space and translate it to the cursor position in the screen space. After mouse cursor is navigated to the desired location, mouse button clicks are triggered by voice recognition of particular commands like right, left, double, drag, drop for right click, left click, double click, drag and drop a file respectively.

Experimentally Head position is Captured by a Camera mounted on top of the monitor, a continuous video is framed into images where image consist of user's head on changing direction and translated it to the cursor position in the screen space. The main contributions and advantages of our work are summarized as follows. 1) We proposed a new camera mouse system [2] where a hands-free system is proposed which is co-ordinate with the user's interaction. 2) A low cost head pose and motion estimation scheme is introduced with combine approach of voice recognition. 3) We designed an efficient strategy for virtual mouse manipulation in real-time. Haar-like feature is one of popular real-time algorithms that can detect objects correctly [4][5]. The system consists of a robust real-time vision based head tracker, a head position and motion estimator, and a virtual mouse control module. Cursor position is navigated and fine-tuned by calculating the relative position of tracking window in image

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

space and the user's head by use of vision based algorithm.

II. SYSTEM DEFINITION AND FRAMEWORK

Implemented on typical PC, our camera operated mouse provides a specific virtual human interface for hand-free mouse control. Fig.1 illustrates the system framework of camera mouse.

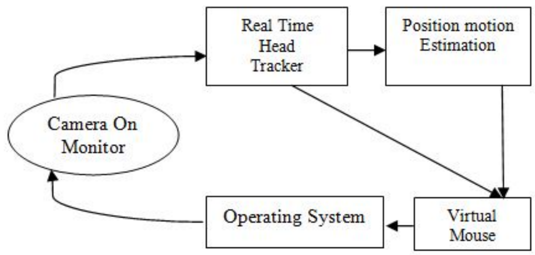

Fig.1 System definition of vision based mouse

By using a common web camera without any calibrated lens, or sensor the system processes each frame of the captured video in real time. The above figure is given by Yun Fu in "hMouse: Head Tracking Driven Virtual Computer Mouse". User's face/head is first automatically [6] detected and tracked by a robust and reliable head tracker. The head pose and motion measurable characteristics are further estimated by analysing visual sight. With basic synchronization control visual tracking module navigates cursor and control virtual mouse buttons using the received voice commands. Operating system finally responses all mouse events generated by this PUI [6].

III. ROBUST REAL TIME HEAD TRACKER

In each tracking cycle, for a captured video frame, camera operated mouse first processes automatic real time Head detection.

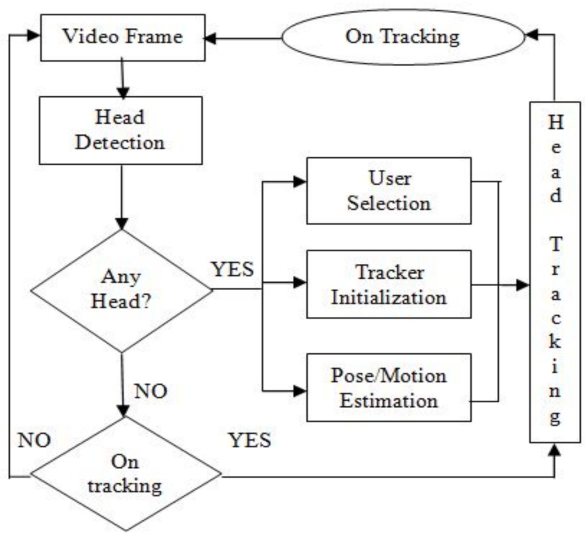

Fig. 2 Framework of Haar base mouse detection

The first detection window determined when the cursor by head track is turned on is then selected as the initialization of the real-time tracking module.

The head tracker enables the user to switch between detection and tracking which is complementary exhibiting very reliable performance. There are many existing pure frontal / profile face detection system that shows robustness against occlusion and illumination variation. If any head is detected three procedures get called i.e. user selection, tracker initialization and head position estimation.

These three procedure are sequential out of which first module is capable detecting user from multiple head depending upon more head exposed area once head is detected tracker is initialized and finally head position estimation is done by vision based algorithms. The detection and tracking are obviously complementary to each other to handle more difficult cases such as user jumping, user come in and out, multiuser occlusion, turning around and large degree rotation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A. Estimation of Head Position and Motion

Proposed Method for Head Movement tracking is Haar cascade object detection method. Haar Cascade method is product of Voila-Jones implementation. Video processing is continuous framing of pixel values, where pixel is smallest unit of process. In proposed work every frame is searched for head position and we highlight this portion by rectangular box.

Normally in case of detection head it is very easy task to detect head because of its geometrical significance. Each person has two eyes, one nose, one mouth and two ears. These are our main four objects to recognise. Other animals have also this object but symmetry in every human face and any other animal face in is different. For every image processing we always convert the image into gray-scale format because it is uniquely converts all variable types images into bright and black image, generally brighter image pixel is considered one and black image pixel is considered zero. If we observe that human face consist of eyes somewhat deeper than the nose and mouth so we take here advantage of Gray-scale image. Gray-scale image of human face consist brighter portion on eyes and blacker portion on remaining sides. Now we apply image integral method on extracted gray image. Image integral is started from left top of image to the right bottom in search of brighter portion in an image. The Haar-like feature is the statistical math model of trained kernel. Then we classify kernel with the database [8][9][10].

This procedure is very tedious to implement for every frame in real time video so as an alternative we make the .xml file having rectangular box co-ordinates in it of the object to track. We call this file for every frame in video and match frame image with the object that is stored in co-ordinated in file.

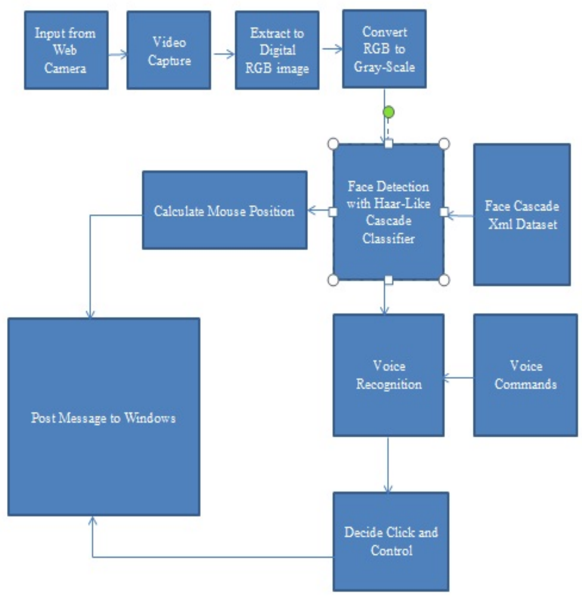

Fig.3 Full diagram of our proposed algorithm

B. Scaling

Based on the estimated object size in (length, width) by Haar detection we rescale the window size as ((length + width) /2). Then we determine scaling factor to set sensitivity of mouse motion on windows screen.

C. Horizontal and Vertical Motion

The horizontal and vertical motion is calculated differently. In detection mode, the reference motion point is
$R_m = ((W_{left} + W_{right}) / 2, (W_{top} + W_{bottom}) / 2)$.

IV. MOUSE CONTROL

A. Move Cursor in Large Scope

Mouse cursor moves according to the relative mapping between scend and processing windows screen motion. Relative cursor motion is multiple of factor of size of the tracking window. Normally it is six time the tracking window speed.

B. Left and Right Button Click

The event of clicking left and right are handled by voice recognition commands. Voice recognition provides you interactive

Volume 3 Issue III, March 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

mode with computer. User has to say some commands such as Left, Right, Double, Drag and Drop. Only these five commands are allowed to handle clicking. The functionalities of these commands are for clicking left, right, open particular file, dragging some file and dropping according to the above sequence.

There are number of voice recognition system that are available on market, the industry leader are IBM and Dragon system. Voice recognition can refer to one of two types of computer science: forensic voice identification or speech-to-text capability. The two basic types of SAPI engines are text-to-speech (TTS) systems and speech recognizers. TTS systems synthesize text strings and files into spoken audio using synthetic voices. Speech recognizers convert human spoken audio into readable text strings and files. The program matches the audio signature of speech with corresponding entries in the database. The SAPI application programming interface (API) dramatically reduces the code overhead required for an application to use speech recognition and text-to-speech, this makes speech technology more accessible and robust for a wide range of applications. The SAPI API provides a high-level interface between an application and speech engines. SAPI implements all the low-level details needed in control to manage the real-time operations of various speech engines.

V. EXPERIMENTAL RESULT

Our project first finds the location of head in video frame and calculates the co-ordinates from the image and it maps coordinate screen with X and Y axis position, then it detects the voice commands. We have particular set of commands and we match these commands with the inputted commands and we perform right and left commands as follows.

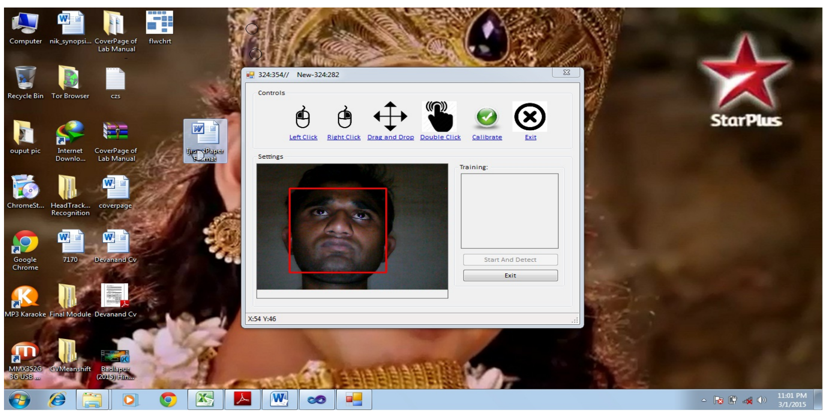

Fig.4 Face is detected and left clicks is done by voice command

We evaluate the system performance by setting up experiments in distance versus sensitivity parts in this we vary distance parameter between user and the built-in computer notebook camera as shown in figure fig.5. The camera is a 0.3-megapixel with built in camera with approximately 9 frame/seconds. And the computer has a 2.27GHz Intel Core I3 Processor with 3 Gigabyte of RAM and Windows 7 OS 32 bit. For Each experimental setup we consider approximately 60 seconds approximately 500 frames. We assume that luminance at user is more or less than luminance at camera. The distance between web camera and user is approximately 45 centimeters with more than 90 percent sensitivity as shown in fig.5.

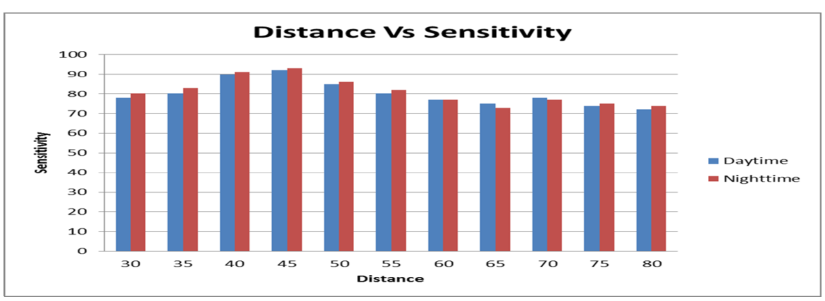

Fig. 5 Experimental result of frontal head pose at various distances.

VI. CONCLUSIONS

We have presented the novel head tracking driven camera mouse system which significant cherished aspect to remove or to

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

diminish manual controllers such as joysticks or keyboards.

This new generation of technology eases the interaction between human and computer eliminating the intermediate devices, and gives more chance to people with disabilities, especially in their hands or arms, to be able to utilize different aspects of technology such as entertainment, during their life. The current version of Mouse is limited by its cursor control mode. We still need to improve the smoothness of cursor navigation algorithm so that users may feel easier to get use.

REFERENCES

[1]. Amer al-rahayfeh and miad faezipour “Eye Tracking and Head Movement Detection: A State-of-Art Survey” IEEE trans. on Rehabilitation Devices and System, 2013.

[2]. M. Betke, J. Gips, and P. Fleming, “The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People with Severe Disabilities,” IEEE Trans. on NSRE, Vol. 10, No. 1, pp: 1-10, 2002.

[3]. T. Morris, and V. Chauhan, “Facial Feature Tracking for Cursor Control,” Journal of Network and Computer Applications, Vol.29, No.1, 2006, pp. 62-80.

[4]. V. Srisuwarat and S. Pumrin, "A Vision Module Algorithm for Human Face Detection," in The Proceedings of the 31th Electrical Engineering Conference (EECON31), 2008, pp. 1165-1168.

[5]. Yun Fu and Thomas S. Huang “hMouse: Head Tracking Driven Virtual Computer Mouse” IEEE Workshop on computer vision, 2005.

[6]. H.Lue, R.Zhang “ Head Detection and Tracking by Mean-shift and Kalman Filter” Innovative computing, Information ICICIC, 2008.

[7]. L. Jian-Zheng and Z. Zheng, ``Head Movement Recognition based on LK algorithm and Gentleboost," in Proc. IEEE 7th Int. Conf. Netw. Comput. Adv. Inf. Manag., Jun. 2011, pp. 232-236.

[8]. C.Kraichan, S. Pumreen “Face and Eye tracking for controlling Computer function” IEEE conf. on Ele. Electronics, Information, 2014.

[9]V. Srisuwarat and S. Pumrin, "A Vision Module Algorithm for Human Face Detection," in The Proceedings of the 31th Electrical Engineering Conference (EECON31), 2008, pp. 1165-1168

[10] O.M. Vultur and O. Gherman, "Human Arm Detection using Haar-Like features," DOCT-US, an. II, 2010.