The Nucleus **50**, No. 4 (2013) 369-379

The Nucleus
A Quarterly Scientific Journal of Pakistan Atomic Energy Commission
NCLEAM, ISSN 0029-5698

A REVIEW OF VISION-BASED PERCEPTIVE INTERFACES: MOUSE ALTERNATIVE APPROACHES

U. TARIQ, M.D.A. ASIF, *M.H. YOUSAF and F. HUSSAIN

Department of Computer Engineering, University of Engineering and Technology, Taxila, Pakistan

(Received July 22, 2013 and accepted in revised form October 31, 2013)

Graphical User Interfaces based on WIMP (Window, Icon, Menu, Pointer) paradigm has been a popular choice of interfaces for many systems and has made Human-Computer interaction simpler and easier. In the past few decades great research efforts have been made to make user friendly Human-Computer Interfaces. The Perceptual User Interfaces aims to focus this new form of HCI (Human Computer Interaction). In this study we describe different Vision Based Perceptive Interfaces that make use of different computer vision techniques for tracking and gesture recognition that may serve as an alternative for outdated mouse. The major focus of the work presented here is on the critical evaluation of different vision-Based interfaces that will give the perceptual ability of vision to computers and will make human-Computer interaction more human friendly interaction. The comparative study and critical evaluation using Human Computer Interaction approaches highlights the strength and weakness of the existing vision-Based interfaces. The emphasis of our research is focused on summarizing the literature by highlighting its limitations and strengths.

Keywords: Computer vision, Perceptual user interfaces, Vision-based interfaces, Gesture recognition, Feature tracking

1. Introduction

Mouse provides an effective and fast means of communication with the GUI (Graphical User Interfaces) based on WIMP (Windows, Icon, Menu, Pointing devices) paradigm. GUI's have been a popular choice of interface for many systems like Window's, Linux, and MAC OS etc. As the way we use computer changes, GUI's will not be able to meet user need of interaction. In order to include wider range of scenarios, user requirements and preferences, interfaces are required to be more adaptive, unobtrusive and natural. Perceptual interfaces (for more details read section 2) focus the area of Human Computer Interaction that makes computer interaction more human-human like. Providing Computer with perceptual abilities makes computer interaction more interesting and human like. Perceptive interfaces provide computers perceptual abilities of vision and speech understanding etc. Vision-Based perceptive interfaces add the perceptual ability of vision to computer and can control the cursor and simulate button presses/clicks using computer vision techniques of tracking and gesture recognition.

The paper is divided into five sections. In Section II we will discuss vision-Based interfaces that may replace traditional mouse. Section III will present the alternative mouse that exists. Section IV will emphasize the strong and weak points of vision-Based interfaces Section V will present future work and conclusion.

2. Computer Vision Based Human Computer Interaction

Most of modern graphical user interfaces are based on the WIMP (Window, Icon, Menu and Pointing device) model [1]. WIMP based interfaces employ a pointing device to communicate with the computer (e.g., trackball, joystick, mouse etc.) to select icons available on the display screen which performs a predefined action and commands from the drop down menus. WIMP paradigm requires user to have physical contact with the pointing device to convey the user input. People with severe disabilities and limited motor functions find it difficult to use such interfaces also if the users both hands are busy then interaction with the computer becomes challenging. Over time the computer hardware has become powerful and inexpensive and advancement in the field of computer vision, designing vision-based Perceptual User interfaces (PUI) has been the focus of research for the past few years [2, 3]. These systems make use of video camera to detect visual cues of user, such as head motion or motion of eyes to control the computer cursor/pointer and gesture recognition to generate mouse button presses/clicks.

* Corresponding author : haroon.yousaf@uettaxila.edu.pk

M. Turk and G. Robertson [4] explain the perceptive user interfaces that add perceptual capabilities of human to computer giving awareness to computers to understand face and hands gestures and user audio input [5]. Human communication is multimodal; we use multiple modalities when communicating face to face. The aim of perceptual user interfaces is to design interactive systems modeled on human-human interaction allowing people to communicate with the technology in a similar way as to how we interact with each other and the world e.g. WIMP based GUI's augmented with vision based perceptual interfaces [3]. Thus interfaces based on computer vision are being exploited more and more as a new or alternative input modality in the class of perceptive user interfaces. We will focus here only on the use of head, eye, nose and hand/finger tracking to control the pointer in WIMP paradigm and on facial expression recognition, gesture recognition and blink detection to generate mouse events.

2.1. Nose Tracking VBI (Vision Based Interface)

Human nose is an important feature when it comes to facial tracking. It is clearly visible when head is moved. Perceptive VBI are focused at building hand-free alternatives to mouse, joystick etc. The nose tip can be captured and tracked by such system using video camera or a web-cam. The movement of the nose can be translated into cursor movement. Thus nose can act as pointing device, nose being located in the middle of the face is an easier feature to move the pointer.

D. Gorodnichy [6] discusses the nose importance for tracking faces. Other facial features look different when head is moved or rotated. The convex shape of the nose makes the tip of the nose easier to track because characteristic pixel pattern can be seen even when nose it rotated or tilted. In his work Local template matching technique explained [7] is used for tracking. The technique scans the window of interest through a peephole mask then after obtaining the feature vector it is compared with the template vector. The peephole mask and the template vector is known in advance. Local and global approaches for face tracking either lack precision or robustness to head motion (rotation and scaling). Global approaches lack precision where as local approaches suffer from head motion robustness. Selection of features such as eyes, nostrils, mouth etc. is also another concern for local approaches. Use of multiple local features suffers from the jitter problem of determination of the pose. Using single feature solves the jitter problem and a proper design leads to robust and precise tracking. Construction of a vector feature template that stays invariant when user is moving solves the problem related to robust tracking. The point on the nose surface that is closest to the camera is referred to as nose feature. A facial feature that is easy to track has the property of uniqueness (feature vector and other vector should be as far apart as possible in multidimensional space) robustness (feature vector does not vary as the tracking proceeds) and continuity. It is shown through experiments that nose feature is both robust to rotation and is precise.

A system called Camera Mouse originally developed by Margrit Betke, James Gips and Peter Fleming [8] works as a replacement of mouse for windows computers. In their work the system tracks the movement of the user which are then translated into pointer movement on the screen. The original idea involved two computers that linked together, the vision and user computer.

The vision computer runs the visual tracking algorithm and forwards the tracked position of the feature to user computer. The functionalities of the two computers can be incorporated into a single computer. Different body features such as nose, lips, eyes finger etc. can be used for tracking. Initially user selects the feature to be tracked by clicking the feature. The sub-image then serves as a template to find the position of the feature in the next frame. A search window centered at the position of the previously found feature is searched by shifting and correlating the template with the other sub-image in the search window. The sub-image with the highest correlation coefficient is found. The coordinates of its center are interpreted as the coordinates for the cursor and the template is updated with newly found feature in the current frame.

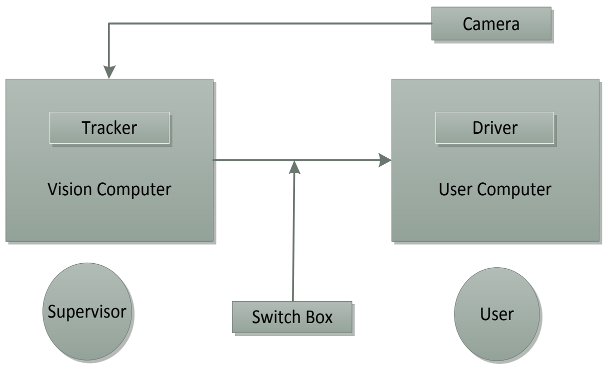

Figure 1. The camera mouse system (Courtesy Betke et al. [8]).

The user computer runs the special driver program based on eagle eye system [9] that runs in the background, it scales the x, y coordinates to current screen resolution. The driver program provides option for mouse click using “dwell time”. The driver program generates the clicks for the user with the “dwell time” option if the pointer stays for 0.5s within 30 pixels of the target.

Camera Mouse software can be downloaded free from [10]. The software works well with Logitech HD pro web-cam C910 [11] or with Microsoft life cam studio HD [12], with some low resolution cameras the jitter problem makes it difficult to control the cursor. For left click and double click dwell time and double click option need to be selected, which requires supervision for people with severe disabilities. It is annoying and at times difficult to constantly select the double click button in camera mouse window. The system is not completely automatic as it requires some form of initialization (selection options/ feature selection).

A system proposed in ref. [13] that uses tip of the nose as a pointing device and human eyes blink feature is used to simulate mouse events. The design of the system is broken down into three modules

1. Tracking of the facial features.
2. Integration of nose tip with the mouse.
3. Using eye blinks as mouse events.

Face is detected in real time using distance information, six segmented rectangular filter and template matching technique. The SSR filter which is a rectangle is divided into six segment uses the bright-dark relation around the area of between eyes. “Integral Image” mentioned [14] is used to calculate pixel value sum in every segment of SSR filter. The entire image is scanned with the SSR and for each segment gray level average is calculated from integral image, bright-dark relation among the segments is used to find out if the center is a candidate for between eyes. Stereo camera is used for calculation of suitable template size and distance information. A technique called average between the eyes is used to find true candidates for between eyes. Once the face is detected it serves as a region on interest to find nose tip and tracking of nose tip using template matching inside ROI. Finally in the ROI blink is detected only when the cursor is at stand still.

In ref. [15] a VBI interface using standard USB web-cam based on nose tracking is proposed for hands-Free computer accessibility. The proposed system makes use of combination of many computer vision techniques for gesture recognition, feature detection and tracking. The system is composed of two primary modules Initialization and processing. Prior to initialization system require head of user positioned such that is faces camera without any orientation. Figure 2 shows the general over view of system main modules.

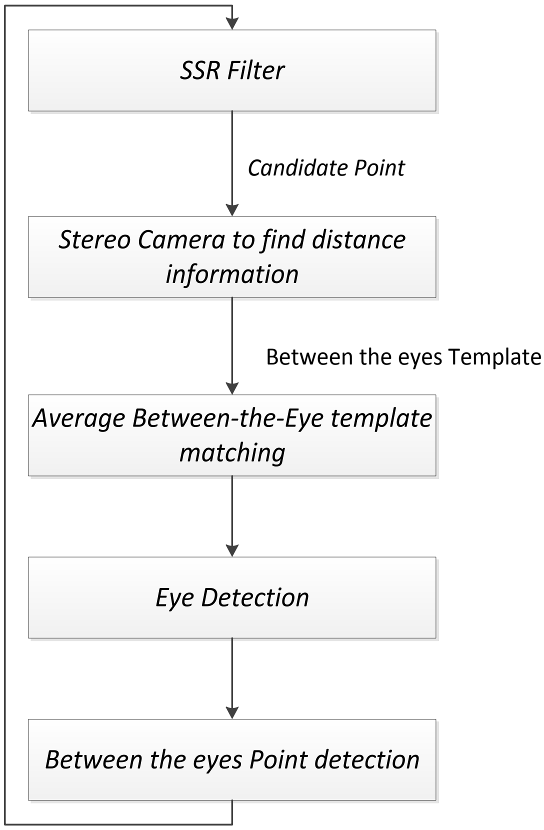

Figure 2. Flow diagram of real time face detection (Courtesy Sumathi et al. [13]).

Once system initialization is completed it works fine for head orientation as well. The purpose of initialization module is to extract facial feature models. It locates the user's face also learns skin color and detects location of initial facial features. It also extracts properties of image like appearance and color. The entire process is completely automatic. Nose is used for head tracking because of its symmetry and visibility. Gesture recognition module is responsible for recognizing left right wink. Mouse events like left right and double left click are available using wait-until click over graphical keyboard.

Chuhan and Morris [16] in their work proposed a cursor control system, using nostril positions relative to face to control the cursors position. The video data captured using a camera mounted below the monitor. Face is detected by initially applying color filter to the video data. Application of filters results in multiple blobs or skin regions. Reliable statistical methods are used to detect the largest blob or largest skin region which is assumed to be the face of the user. Using heuristic rules nostrils are located.

Another System proposed [17] makes use of nostrils for tracking. The system proposed is founded on novel template matching technique [18] that solves the problem of occlusion. The proposed technique is also robust to rotation and scale variation. The template matching technique compares gray level intensities of a chosen template to a gray level image that makes the region of interest. For comparison purpose a similarity metric is used, here normalized sum of squared differences (NSSD) is used. To solve the problem of occlusion and rotation, original template as well as the template obtained from previous frame is used. Two similarity matrices resulting from the comparison are scaled using a weighting factor. These matrices are added up and the location that has the smaller result is selected. That is the best match and is used as a template for the next frame. To make the algorithm computationally fast, Kalman filter is used to narrow down the search space. To generate mouse events nodding and face shaking is used (nodding left click, face shaking for right click). There are certain issues with the system, it requires initialization procedure. Camera position needs to be adjusted so that user face is in the center of view. The feature to track in this case nostril is needs to be selected that will serve as a first template.

2.2. Head Tracking Vision-Based Interfaces

In the literature the concept of head tracking as an interactive system input component has been around for some time. Head tracking, is similar to other facial feature like nose tip, nostrils, eye etc., may serve as an alternative for cursor control in vision based perceptive interface.

Gaver [19] has presented virtual window system that uses movement of head to control the movement of camera in remote office, this allows the exploration of the scene rather than a screen which is showing moving pictures of head movement. The prototype that was built was noisy, slow and inaccurate and could not be used for a long time.

Toyama [20] suggests use of head tracking to control the pointer. A hands-Free cursor control is provided by moving the cursor in the direction of the nose. The incremental focus of attention [21] structure robustly tracks 3D face orientation and position. Incremental Focus of Attention (IFA) for robust tracking was found on several cues like intensity template and color that collaborate to tracking under inimical visual circumstances. Recovered pose from tracking is used to calculate the crossing between the flat surface of the monitor and assumed ray elongated from user nose.

Francois Berard [22] has presented a system called the “Perceptual Window”, perceptual in a way that it reacts to user action rather than direct physical contact. The interaction technique presented uses the head motion to control window view-point location. A non-intrusive tracker is used to track over time. The technique called correlation method [23] is used for tracking. The tracker outputs the (x, y) coordinates of the target, that are combined over time to give the translation of target. The target here is the head, in this way system is informed of the face motion. The tracker output is used in both position control interaction (moving object with in the document) and rate control interaction (navigating inside the document). The system focuses on improving GUI interaction. It augments the standard interaction rather than completely replace it.

A prototype system presented [24] proposes on controlling of PC by means of head movement and voice control. The system is multi-modal and the modes of communication are gesture and speech. The system if developed with an intent to provide people with severe disabilities to interact with the system. The feature used here for tracking is head/face. The head movements are translated into mouse pointer movements and voice is use for button presses. The system requires initialization process so the face can first be detected using skin color information and later tracked. Commands for right, left, double clicks and others are available through voice mode interaction.

A relative virtual mouse is proposed [25] based on the comprehension of head movements via low cost camera. The device is designed for the people with mobility impairments. The image processing algorithm used for the motion detection is based on the [26] (Optical Flow or image motion estimation), that compares two consecutive images. The algorithm used for the implementation of the virtual mouse consists of

- User Detection
- Head movement detection
- Eye blink detection
- Open mouth detection

User detection involves detection of head, that requires the user to be in front of the camera and the head of the user in central part of the image. The eye blink can be detected and can be used as a mouse button push or clicks. Normally the natural eye blink is very fast and is hard to be picked up by the camera. The unnatural eye blink can serve as an alternative for mouse clicks. Motion detection algorithm implemented here makes open mouth detection easy. Open mouth can also be used for button presses/clicks.

"hMouse" a novel head tracking based mouse is presented [27]. The hMouse is composed of a real time robust head tracker, a virtual mouse control module and head motion estimator. The hMouse tracker makes use of 2D tracking complementary switching approach with an interactive loop. The hMouse calculates users head, yaw, tilt, roll, and scaling, vertical and horizontal motion for the control of the mouse, from trustworthy tracking result. In the image space the relative position of tracking window is calculated that is translated to position of the cursor in the screen space. For fine tuning tilt of head and yaw circumvolution are used. When the cursor has moved to the location desired the head roll rotation triggers the mouse events. "hMouse" therefore has a potential for being used as a hand-Free alternative for mouse. The applications of hMouse extend from computer games, robot control, and machine guidance to people with disabilities and elders. The structure of the hMouse is shown in Figure 3.

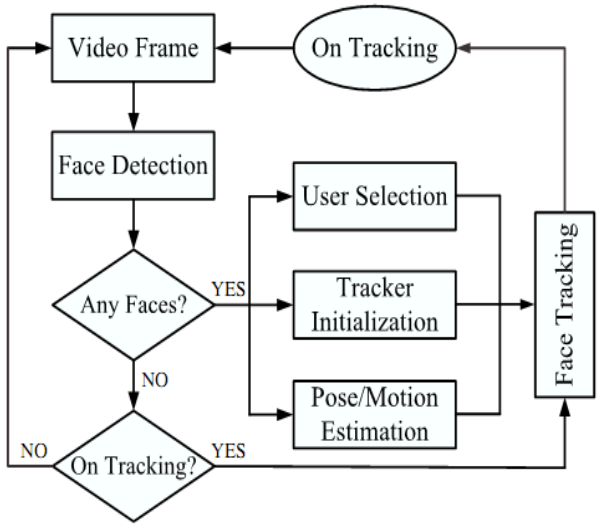

Figure 3. Structure of Head Tracking Mouse Tracker (Courtesy Fu Y. et al. [27]).

2.3. Hand / Finger Tracking Vision-Based Interfaces

Hand/Finger tracking provides another means for cursor control. There are many situations where hand/finger tracking based interaction is more intuitive for example in a meeting taking place in a room with large display where several people are discussing their ideas by drawing pictures or controlling a remote large display in a smart room.

A Vision-Based method to implement virtual mouse is proposed in [28]. An efficient finger tracking algorithm and set of messages for controlling windows based application are presented. The application structure based on virtual mouse has two parts.

1. Virtual mouse server.
2. Application based on virtual mouse (client).

Server is responsible for tracking finger tips, constructing and sending appropriate messages to client. Client responds to these messages by calling API functions that are provided by the server. User is required to wear Finger-Knots of a given color for fingertip detection and tracking purposes. To detect the fingers, Video image is transformed from RGB to HSV color space, back projection on color histogram is used to build probabilistic distribution image. Binary image of probabilistic image is obtained via binarization. Morphological opening and closing [29] are applied to binary image to remove background color and holes in the image. Edges of the regions are obtained and each region is labeled using minimum bounded rectangle. Having found the finger tips the centroid is calculated to simulate pointer. Different messages defined for the mouse events and the motion of the fingertip in and out of the shooting area. Mouse operation of clicking/button press is simulated using two fingers and actions are recognized by BP neural network.

In their research on virtual mouse [30] Xing Feng and Kai Huai have presented a virtual mouse that allows six-degree-freedom. The virtual mouse is founded on hand gestures/expressions using a web-cam for video input. Hand is tracked using CAMSHIFT (Continuously Adaptive Mean-shift) algorithm according to color of the skin and to recognize poses finger gestures are used. The virtual mouse uses easy gestures therefore the mouse can be used by people with disabilities. Figure 4 shows system diagram.

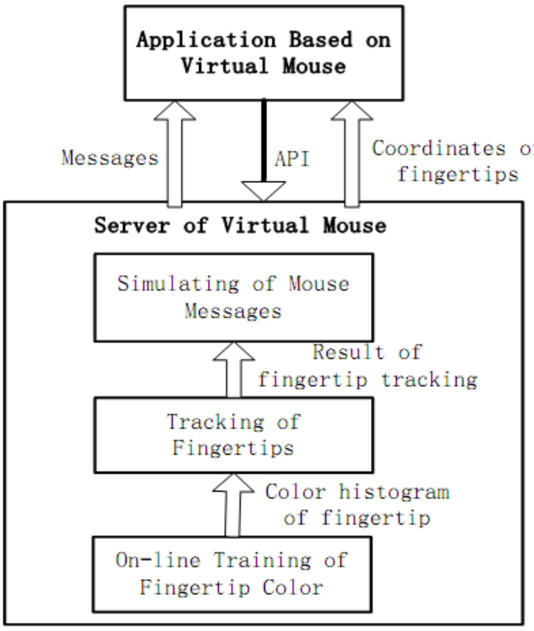

Figure 4. Application structure based on virtual mouse (Courtesy Wensheng et. al. [28]).

In ref. [31] a virtual mouse that is based on the tracking of the finger movements is proposed (Figure 5). The system presents an idea to control the computer or TV etc by means of finger movements without physical contact. A fast robust method to track the finger tips is proposed. Using segmentation the pixels of the skin-color is detected by means of chrominance component of the input received from Complementary metal oxide semiconductor (CMOS) sensor. In order to reinforce the pixels of the skin-color density regularization is used. To minimize computational cost an effective technique of window search is used. Tracking of fingertip is used to find the tip of finger position. Also the clicking is provided using specific movements. The virtual mouse system is implemented on embedded Linux.

3. Alternative Mice

Some other types of mice which combine multi touch sensing with the standard capabilities of virtual mice are described in this section. The goal is to change normal interaction of desktop pointing with touch and gestures. Below is the description of few alternative mice.

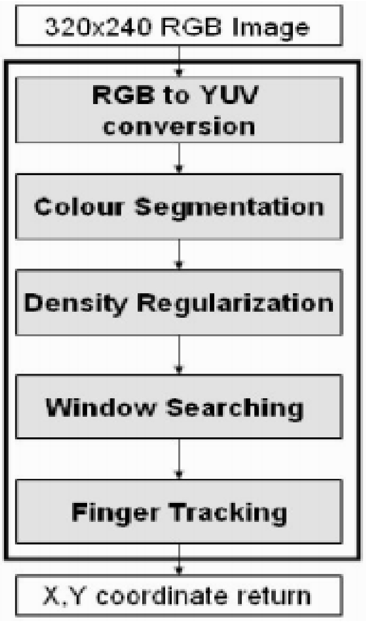

Figure 5. YUV color space finger tracking flow (Courtesy Wah W. et al. [31]).

FTIR Mouse: Frustrated Total Internal Reflection is a multi-touch mouse design which enables multi-touch input on interactive surfaces. According to this idea a thin sheet of acrylic is lit on the edges with infrared light. When the finger is pushed on the sheet of the acrylic, it causes Infrared light to scatter away from the finger, this can be detected using an infrared camera which is capturing the surface.

When light reflects from interface to lower index of refraction medium (like from glass to air), the light becomes refracted which depends on the angle of incidence and after a certain angle of incidence, it follows a phenomenon which is total internal reflection (TIR).This basic phenomenon is used in FTIR mouse [32]. The edges of acrylic sheet should be polished clear for optical waveguide. Common glass surface for this mouse cannot be used for poor luminance. Image processing operations like rectification, connected component analysis, noise removal, and background subtraction are required for every frame that camera captures, whereas computer vision techniques are used to convert the whole sequence into events of interaction through this mouse. FTIR mouse gives a binary signal of user touching the surface which makes mouse interaction vigorous in operation. The FTIR technique does have limitations as sensing of fingers is restricted to camera field of view and secondly IR sensing also have restrictions in its usage in sunlight as they do not work properly in sunlight.

Head Mouse Extreme [33] can serve as a replacement for mouse for people with limited abilities to move hands. It translates the user head movement into computer pointer movements. As the head of the user moves so is the pointer on the screen. It connects to the computer through as USB port and no special software is required by the Head Mouse. It receives power from host device over USB port.

***ORB Mous*e:** It provides a multi-touch sensing on its semi-circular surface by combining an Infrared camera and a source of Infrared illumination. ORB mouse is different from FTIR Mouse, because the illumination is not totally internally reflected inside the shell of the mouse, rather it spread outwards from the center of the device to the edges of the shell, and then reflected back into the camera by user hands that come into close to the semi-circular surface of the mouse.

The basic principle of operating ORB mouse is same as interacting with certain interactive surface technologies that use diffused IR illumination [34]. The components of orb mouse includes optical sensors, Camera, mirror, IR LED's for illumination purpose, mouse buttons and diffused surface. The camera inside the ORB mouse is directed towards the internal mirror of mouse which provides very wide angle view of mouse surface instead of pointing camera towards the surface. The semi-circular shape of Orb Mouse is anticipated for easy gripping and the curves of mouse are made such that user's fingers experience a smooth shell while moving from one side to the other and from front to back. The communication area of ORB mouse is larger than FTIR mouse which allows whole hand and fingers in interaction.

ORB mouse has also some drawbacks. It is also sensitive to light as FTIR mouse. The reflected IR images are much more distorted by ORB mouse as images are less distorted in FTIR mouse. Some techniques are used to overcome the distortion as mentioned [35]. Images are very low contrast by camera of ORB mouse.

Ion Wireless Air Mouse Glove [36]: The wireless air mouse glove can be used as an alternative for the mouse for computer interaction. The glove has button available on the index finger for left, right clicks and are pressed using the thumb. To scroll the glove wearer uses movements of wrist to go up and down. The mouse is wireless and can be charged with a USB port. From 35 m away the user can control the computer.

Side Mouse detects the user fingers when they touch the table surface, rather than a mouse. Same design is previously implemented on mobile phone with the help of IR sensors placed on the sides of mobile phone [37]. In case of the side mouse, user fingers should touch the surface of the table right in front the side mouse while the mouse is placed under the palms of user. The front end of the mouse contains forward sensing camera which is mounted behind Infrared pass filter. This camera includes wide angle lens so that when user places fingers in front of it, the mouse detects it with the help of IR sensors. When fingers are placed on surface table, IR rays are reflected back to the camera which detects the position of cursor with the help of micro-controller.

***Arty Mouse*:** Arty mouse is short form of articulated mouse which is basically based on the concept of side mouse. In this type of mouse, the user palm rests on the base of the unit and two articulated arms attached to mouse extend that can be moved without restraint on the table with thumb and first finger. The design of this mouse takes three optical mouse sensors, one under base of the mouse and one under each articulated arm of the arty mouse to detect the proper position and placement of cursor. The base of the arty mouse contains similar circuit as Bluetooth optical mouse and makes it wireless. The components of this mouse are also discussed [38]. The two articulated arms are placed on the mouse using pivots. Each arm also contains a button which is same as in normal mouse have two buttons. One of the key advantages of this arty mouse is that it tolerates high resolution optical mouse sensor which is located under two user fingers which gives accurate results as compared to other mice discussed above.

EvoMouse [39]: *The m*ouse is evolution of traditional mouse. With EvoMouse the finger is used as a pointer and requires no physical pushing of the mouse. It works on any flat surface. Mouse operations are performed using fingers only. Hand gestures are used to control right left clicks and select etc.

***Finger Mouse [40]*:** It is another alternative for the traditional mouse ideal for use at home, offices and for the mobile travelers. Works by strapping it to the index finger and operates on any surface.

Table 1. Pros and Cons of pre-described interfaces based on computer vision techniques.

No.	First Author Ref.	Cursor Control	Mouse Events Generation	Pros.	Cons.
1	*K. Toyama [20]*	*Face/ Head Tracking*	*Not Implemented*	*Alternative for hand-free cursor control, User friendly interface.*	*Dependence on preselected feature like eyes and lips needs to be eliminated, unwanted jitter in tracking pose due to independent movement of features relative to face, No method to simulate mouse clicks/button presses*
2	F. Berard [22]	Face/ Head Tracking	Not Implemented	Can perform limited navigation task, Complements standard interaction	Tracking fails if target is too fast, Tracking fails if target goes through substantial circumvolution or large distance change, Tracking fails if neighbor-hood target is similar to target
3	A. Ismail [24]	Face/ Head Tracking	using Voice Commands	Allows multi-mode interaction, illustrates accuracy and speed for real time applications, A possible alternative for complete hands-Free interaction	Drop in performance in the presence of dark background, robust voice control is an issue.
4	T. Pallejà [25]	Face/ Head Tracking	Eye blink or mouth	Hands-Free alternative for mouse, low CPU use, Ease of computer accessibility for people with disabilities, Free application available at [45], A possible alternative for hands-Free interaction, Head detected automatically	Slow pointer speed, Not compatible with all web-cams
5	Y. Fu [27]	Face/ Head Tracking	Head Roll Rotation	Tracks single head from multiple heads, Ease of access to computer for people with disability, Works under extreme movement, jumping and large rotation that makes it good candidate for games.	Head pose estimation module not accurate for judgment of user intention, Further work needed to improve accuracy.
6	L.I El-Afifi [17]	Nostril Tracking	Nodding and Face Shaking (Left click and Right Click respectively)	Prone to temporary occlusion, Processing speed fast	Not completely automatic requires initialization; Re-Initialization needed if feature is occluded for sufficient time, Tracking of feature fails if head moves very fast.
7	L. Wensheng [28]	Finger Tip Tracking	Two Fingers (Action recognized by BP neural Network)	Good Performance Under cluttered background	Requires user to wear finger-Knot of given color.
8	V. Chauhan [16]	Nostril Tracking	Not Implemented	Computationally inexpensive algorithm, Works well with low resolution Web cam, Objects of different colors can be tracked by initializing color filter with range and average values at run time.	Cannot be used as an alternative choice for mouse, System gets confused due to similar objects that may exist in the background
9	J. Varona [15]	Nose Tracking	Wait-until-Click over graphical event keyboard	Can be used as a replacement for standard mouse, Already used by people with disabilities	Requires initialization at the beginning of where head has to face screen without any orientation, reinitialization may be needed if feature is lost.
10	W. Wah [31]	Finger-Tip Tracking	Using Finger Motion (still under improvement)	Low cost method as it uses CMOS image sensor, allows mouse to be standalone hardware, low processing power application processor.	Search window under optimization (large search window fast movement can be detected but computationally expensive, small search window fast computation but tracking difficult), Recognition still under improvement, noise effects tracking, similar color object in background effects tracking
11	M. Betke [8]	Nose/Eyes/ Finger/ Head tracking etc. (Any Feature that user clicks on)	Generated by the driver based on Dwell time (Wait of 0.5 second with in 30 pixels of target)	Since the user selects the template multiple similar objects within the near vicinity make no difference on feature tracking, Ease of computer accessibility for people with disabilities, Feature to be tracked is selected by the user, No user borne accessories needed for tracking of features.	Automatic detection of feature not available, user with disabilities require assistance, Some form of initialization required not completely automatic (selection of feature to be tracked, selection of dwell time option etc), cannot totally replace mouse but can be used with the mouse as an alternative, Feature is lost if head shaken very fast.
12	N.A [10]	Nose/Eyes/ Head tracking etc. (Any Feature that user clicks on face)	Left click and double click (available if dwell time and double click option checked in setting)	Since the user selects the template, multiple similar objects within the near vicinity make no difference on feature tracking, Ease of computer accessibility for people with disabilities, Feature to be tracked is selected by the user, No user borne accessories needed for tracking of features, Free application download available from internet, Two computer functionality incorporated into single computer	User with disabilities require assistance, Some form of initialization required not completely automatic (selection of feature to be tracked, selection of dwell time option etc.), cannot totally replace mouse but can be used with the mouse as an alternative, Feature is lost if head shaken very fast, feature lost if face is occluded, Under extreme lighting conditions tracking of feature is very difficult that makes cursor control very difficult
13	S.Sumathi [13]	Nose tip tracking	Left/Right eye blink (Left/Right click)	The system replaces the mouse with human face, Compatible with inexpensive USB CAM,	Requires stereo Camera to find distance information, Tracking of features lost if head shaken very fast, Feature lost if occluded for substantial amount of time.

Cap Mouse: It is basically a capacitive mouse. The location of fingers on the mouse is tracked by capacitive touch sensing mechanism, which is in contrast to capacitive sensing techniques which detect clicks only [41]. This archetype uses a stretchy matrix of capacitive sensors of electrodes to track the location of the user fingers. 1D scrolling [42], or the single finger position [43], the Cap Mouse design includes an accurate multi touch sensor and it is capable to track the locations of all of the user's fingers on the surface of the mouse.

Air Mouse [44]: The mouse provides a stress free interaction with computer unlike traditional mouse increasing the movement of mouse speed and accuracy. It aligns with the ligaments of the hands and wrist. The optical tracking laser , left and right click buttons are positioned adjacent to finger tips and palms so that the user can use both air mouse and touch type alternatively without removing the mouse. It is rechargeable wireless device that has the operational run time of one week.

4. Critical Evaluation

This research work was conducted because with the passage of time and invent of new technologies; researchers and scientist are focused to provide natural-like user interfaces for interaction with computers and machines. WIMP model has served as a most popular choice of many graphical user interfaces for years now. But over the years many research efforts have been made to evolve this model. This paper was aimed to discuss some vision based perceptive interfaces those try to control cursor and generate mouse events using tracking and gesture recognition techniques. Most of the systems control the cursor by tracking some feature like eye, nose finger etc. and generate mouse events either by gesture recognition or by some alternative means such as hold and wait.

A computer mouse works almost under any conditions. However, vision based perceptive interfaces are limited by the computer vision techniques that they employ. In order to augment the WIMP model and use computer vision as an alternative for cursor control requires computer vision techniques of tracking and gesture recognition to be effective under wide variations in conditions (lighting, pose, shadow etc.). Table 1 enlists pros and cons of systems discussed in the section 2 and 3 from the perspective of computer vision techniques used by them. Critical evaluation has been made based on the cursor control, mouse event generation and different parameters for the evaluation of human computer interfaces. Cursor control shows the type of gesture used to operate the device/method. Research work by various authors was evaluated on parameters e.g. technology, ease of deployment, environment, cost-effectiveness, operating constraints, ergonomics and user friendly. The pros and cons enlisted in the table are provided by the authors of the papers.

5 . Conclusion and Future Work

WIMP based interfaces has served as an effective means of human-Computer interactive communication. In this study we have discussed alternative for traditional mouse devices. Vision-Based interfaces described in section 2 can serve as an alternative for traditional mouse.

We conclude that vision-based interfaces are as fast, robust and accurate as compared to computer vision techniques that are used for tracking and gesture recognition. In future with the improvements of computer-vision techniques, these interfaces will serve as effective means of human-computer communication and augment the GUI's to provide more human like human-Computer communication.

References

[1] A.V. Dam POST-WIMP User Interfaces. Communications of the ACM **40**, No. 2 (1997) 63.

[2] D. Gorodnichy, Perceptual cursor - a solution to the broken loop problem in vision-based hands-free computer control devices. In NRC-CNRC Technical Report. NRC/ERB-1133. February 2006. 23 pages. NRC48472, (2006)

[3] M. Turk, Perceptual User Interfaces. In the Proceedings of the 1998 Workshop on Perceptual User Interfaces (1998).

[4] M. Turk and G. Robertson. Perceptual User Interfaces. Communications of the ACM, **43**, No. 3 (2000) 33.

[5] A. Dix, J.E. Finlay, G.D. Abowd and R. Beale. Human-Computer Interaction, Prentice Hall (2003) pp. 365-384.

[6] D. Gorodnichy, On Importance of Nose for Face Tracking, Proceedings of the International Conference on Automatic Face and Gesture Recognition (FG2002). Washington, DC, USA, (2002).

[7] D. Gorodnichy, W. Armstrong and X. Li. Adaptive Logic Networks for Facial Feature Detection, The Proceedings of International Conference on Image Analysis and Processing (1997) pp. 332-339.

[8] M. Betke, J. Gips and P. Fleming, IEEE Transactions on Neural Systems and Rehabilitation Engineering **10**, No. 1(2002) 1.

[9] P. DiMattia, F. X. Curran and J. Gips. An Eye Control Teaching Device for Students Without Language Expressive Capacity EagleEyes. Lampeter, U.K. Edwin Mellen (2001).

[10] CameraMouse, “cameramouse” Internet: http://cameramouse.org/index.html, (Sept. 9, 2012).

[11] LogiTech. “HD Pro Webcam C910” Internet: http://www.logitech.com/en-us/support/ web cams/hd-pro-webcam-c910 (September 9, 2012).

[12] Microsoft Hardware, “LifeCam Studio” Internet:http://www.microsoft.com/hardware/en-us/p/lifecam-studio#overview, (September 9, 2012).

[13] S. Sumathi, S.K. Srivastsa and M.U. Maheswari, Int. J. of Computer Science and Information Security **7**, No. 1 (2010) 147.

[14] P. Viola and M. Jones, Rapid Object Detection Using a Boosted Cascade of Simple Features, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2001 CVRP, 1 (2001) pp. 511-518.

[15] J. Varona, C. Manresa-Yee and F.J. Perales, Journal of Network and Computer Applications **31** (2008) 357.

[16] V. Chauhan and T. Morris, Face and Feature Tracking for Cursor Control, Proceedings of 12th Scandinavian Conference on Image Analysis Bergen (2001).

[17] L. Afifi, M. Karaki and J. Korban, Hands-free'- Interface A Fast and Accurate Tracking Procedure for Real Time Human Computer Interaction, Proceedings of Fourth IEEE International Symposium on Signal Processing and Information Technology (2004) pp. 517-520.

[18] H.T. Nguyen, M. Worring and R.V.D. Boomgaard, Occlusion Robust Adaptive Ttemplate Tracking, Proceedings of IEEE Conference on Computer Vision (2001) pp. 7-14.

[19] W.W. Gaver, G. Smets and K. Overbeeke, A Virtual Window on Media Space. Proceedings of CHI’95 (1995) pp. 257-264.

[20] K. Toyama, Look, Ma–No Hands, HandsFree Cursor Control with Real-Time 3D Face Tracking, Proceedings of the 1998 Workshop on Perceptual User Interfaces, (1998).

[21] K. Toyama and G. Hager, International Journal of Computer Vision **35**, No. 1 (1999) 45-63.

[22] F. Berard, The Perceptual Widow: Head Motion as a new Input Stream, Proceedings of the Seventh IFIP Conference on Human-Computer Interaction (1999) pp. 238-244.

[23] P.A. Anandan, International Journal of Computer Vision **2**, No. 3 (1989) 283.

[24] A. Ismail, A.S. Hajjar and M. Hajjar, The International Journal of Multimedia and its Applications **3**, No. 3 (2011) 15.

[25] T. Pallejà, E. Rubión, M. Teixidó, M. Tresanchez, A.F del Viso and J. Palacín, Simple and Robust Implementation of a Relative Virtual Mouse Controlled by Head Movements, Proceedings of Human System Interactions, Conference on Digital Object Identifier (2008) pp. 221-224.

[26] C. Sun, Image and Vision Computing **20**, No. 13–14 (2002) 981.

[27] Y. Fu and T.S. Huang, hMouse: Head Tracking Driven Virtual Computer Mouse, Proceedings of IEEE Workshop on Applications of Computer Vision (2007) pp. 30-36.

[28] L. Wensheng, D. Chunjian and L. Yi, Implementation of Virtual Mouse Based on Machine Vision, Proceedings of International Conference on Apperceiving Computing and Intelligence Analysis (2010) pp. 367-371.

[29] R. Gonzalez and R.E. Woods, Digital Image Processing 2nd Ed. Prentice Hall (2002) pp. 519-556.

[30] X. Wang and K. Qin, A Six-degree-of-freedom Virtual Mouse based on Hand Gestures, Proceedings of International Conference on Electrical and Control Engineering (2010) pp. 257-260.

[31] W. Wah, M. Tsang and K. Pun, A Finger-Tracking Virtual Mouse Realized in an Embedded System, Proceedings of International Symposium on Intelligent Signal

Processing and Communication Systems (2005) pp. 781-784.

[32] J. Han, Low-Cost Multi-Touch Sensing Through Frustrated Total Internal Reflection. In the Proceedings of 18th Annual ACM Symposium on User Interface Software and Technology (2005) pp. 115-118.

[33] Orion Instruments. “HeadMouse” Internet: http://orin.com/access/headmouse (August 9, 2012).

[34] N. Matsushita and J. Rekimoto, HoloWall: Designing a Finger, Hand, Body and Object Sensitive Wall, Proceedings of the 10th Annual ACM symposium on User interface software and technology (1997) pp. 209-210.

[35] H. Benko, A. Wilson and R. Balakrishnan, Sphere: Multi-Touch Interactions on a Spherical Dis-play. In the Proceedings of 21st Annual ACM Symposium on User Interface Software and Technology (2008) pp. 77-86.

[36] B. Ventures. Ion Wireless Air Mouse Glove, http://www.ionwirelessair mouse. com (August 9, 2012)

[37] A. Butler, S. Izadi and S. Hodges, SideSight: Multi-‘Touch’ Interactions Around Small Devices. In the Proceedings of 21st annual ACM Symposium on User Interface Software and Technology (2008) pp. 201–204.

[38] BT. Newton Peripherals, LLC. MoGo Mouse, http://tinyurl.com/yw4lbv. (12.May, 2009).

[39] Celluon, evoMouse, http://celluon.com/shop_evo_mouse.php (August 10, 2012).

[40] Brando, Finger Mouse, http://usb.brando.com /usb-finger mouse_p00187c037d015.html (August. 10, 2012).

[41] Apple, Mighty Mouse, http://www.apple.com/mightymouse (August 10, 2012).

[42] Logitech, Air Mouse, http://tinyurl.com/3qtor8. (August 10, 2012).

[43] R. Balakrishnan and P. Patel, The PadMouse: Facilitating Selection and Spatial Positioning for the Non-Dominant Hand, Proceedings of ACM conference on Human Factors in computing systems (1998) pp. 9-16.

[44] Air Mouse, AirMouse Internet, http://www.theairmouse.com/tech.php, (August 10, 2012).

[45] T. Pallejà, E. Rubión, M. Teixidó, M. Tresanchez, A. Fernández del Viso, J. Palacín, C. Rebate and J. Palacín. Head Mouse 4.2, http://robotica.udl.cat/(Sept 5, 2012).