V-Touch: Markerless Laser-Based Interactive Surface

C.R. Dikovita[#1], D.P. Abeysooriya[#2]

[#]*Department of Computer Science, International College of Business and Technology*

317A, Galle Road, Colombo 4

[1] rukshan_chanaka@hotmail.com, [2] deloosha@icbtcampus.edu.lk

***Abstract* – Human Computer Interaction (HCI) is one of the major disciplines in the field of computing which defines the way people interact with machines. HCI is a combination of computer science, behavioral science, design and several other areas which determines the best and optimized ways to interact with computers. The way people interact with machines changed dramatically after the 21st century due to the advancements of the computing devices. Computers became more portable and embedded to the environment with the improvement of processing power and miniaturization of devices. New concepts such as ubiquitous computing are also emerged with these advancements of technology. Traditional HCI concepts were failed to meet the requirements of the emerging technologies since most of these new concepts requires natural interactions rather using tangible devices. This paper present a computer vision based interactive surface application for human computer interaction which enable users to perform QWERTY keyboard and mouse interactions on a flat surface. Primarily, this research addresses the existing issues related to gesture based interactive interfaces by presenting a novel approach to perform interactions without stereoscopic vision and wearable markers. The proposing approach present a system (V-Touch) based on Haar-like features with single camera and a laser projection module that allows users to perform traditional keyboard and mouse interactions naturally using low cost hardware setup.**

***Keywords* - Interactive surface, Haar-like features, Augmented Reality, HCI, Virtual Keyboard and Mouse**

I. INTRODUCTION

With the emergence of new technologies such as ambient intelligence and ubiquitous computing, human computer interaction has dramatically changed during last decade. Modern computing devices are more compact and equipped with several embedded technologies. Considering the emerging interaction technologies, interactive virtual surfaces are great potential in human computer interaction which provides natural interaction for augment reality and virtual reality applications [1]. Interactive surfaces are surfaces that enable users to interact with the device using their hands, special pen or other wearable devices. The advantage of interactive surface is that the surface can be any flat surface, for instance table, book, wall or even paper etc. Interactive surfaces can be used to display information and to interact with devices using the surface. Hand tracking and gesture recognition is very flexible to use with interactive surfaces due to its natural interaction without any expensive hardware devices.

Number of research has been already carried out relevant to the particular discipline. Malik and Laszlo's Visual Touchpad [2] demonstrated an interactive surface that uses stereoscopic vision and enables users to interact using bare hands. Sixth sense [3] is another gesture based interactive application which uses a setup with projector and camera. This approach requires wearing color markers at fingers to detect gestures. Some early work of literature revealed on virtual keyboards which use colored markers and head mounted displays [4] and 3D optical ranging [5]. Haar-like features were also used for virtual keyboards and gesture recognition for HCI applications [6], [7]. These applications have some limitations since these approaches solely based on Haar-like features

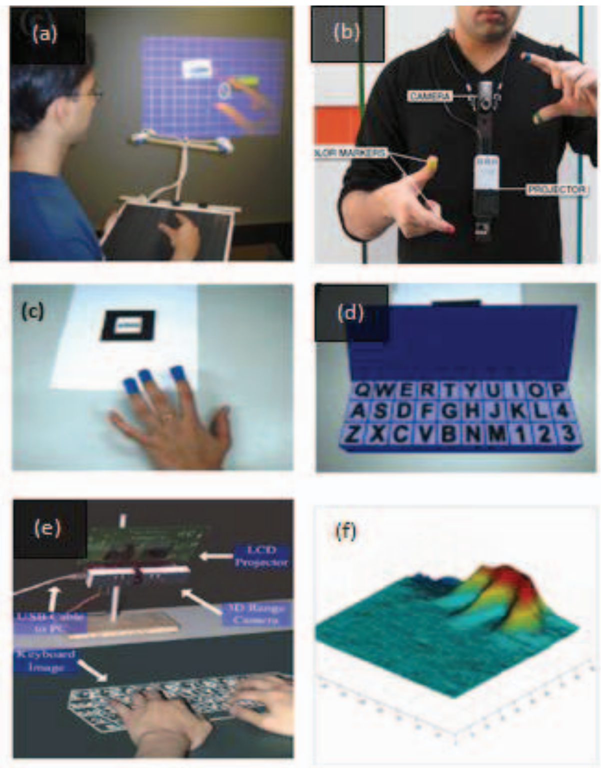

Figure 1: Interactive surface systems: (a) Visual touchpad, (b) sixth sense, (c,d) ARKB, (e,f) 3d true optical ranging keyboard

Recognition of hand postures and gesture can be identified as one of the key challenges in the vision based interactive surfaces. There are number of different hand gesture recognition approaches available including skin color segmentation, deformable templates, Haar-Like features, Hidden Markov Model etc. Haar-like features identified as a robust way to detect the postures, gestures without wearable markers using set of trained classifiers. Viola and Jones [8] presented Haar-like features initially for the detection of faces. This approach was also promising in any object detection and therefore some researches were successful for detection of hand gestures using Haar-like features [9].

As mentioned above, interfaces introduced in recent approaches require either stereoscopic vision or markers to be wear by the user [2]-[5] or they are week in terms of performance and rather difficult to interact [6], [7]. Thus these applications are failed to provide natural markerless interaction with inexpensive hardware setups.

This paper presents an interactive surface to stimulate traditional keyboard and mouse interactions by tracking and identifying hand postures and gestures using single camera and laser based setup with trained Haar cascade to detect hand postures. This study is mainly investigates solutions for flexible, natural and optimized way to interact with computers.

II. HAAR-LIKE FEATURES

Haar-like features are image features that used to object detection. This was originally proposed for the face detection by Viola and Jones [8], [10]. Haar-like features consist two or more adjacent white and black rectangles. The value of the feature is calculated as the different between the sums of the pixel intensity values within the white and black rectangles.

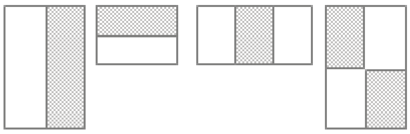

Figure 2: Set of Haar-Like features

Rectangular features of the Haar-like features can be computed using an intermediate representation of an image, which is defined as integral image by Viola and Jones [8]. The integral image at location of (x, y) contains the sum of pixel values above and left of (x, y) inclusive:

$$I(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} i(x',y') \qquad (1)$$

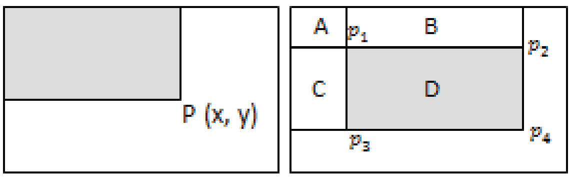

Figure 3: Integral image concept

Sum of the pixel values in the area D can be calculated as:

$P_1 = A$, $P2 = A + B$, $P3 = A + C$, $P4 = A + B + C + D$
$D = P_1 + P_4 - P_2 - P_3$

To detect an object of interest, the image is scanned for particular Haar-like features. Corresponding weak classifier is defined based on the Haar-like feature.

$$h_j(x) = \begin{cases} 1 \ \ if \ p_j f_j(x) < p_j \theta_j \\ \ \ 0 \ otherwise \end{cases} \qquad (2)$$

Considering a particular Haar-like feature f_j, corresponding classifier $h_j(x)$ is determined using equation (2) where x is sub window, θ is threshold and p_j indicate the direction of inequality sign.

Practically it is difficult to detect an object with a high accuracy level when it uses single Haar cascade that defines a single Haar- like feature. Therefore AdaBoost machine learning algorithm is used to improve the accuracy based on series of weak classifiers.

III. DESIGN AND IMPLEMENTATION

This section describes the system architecture of the interactive surface in terms of logical and physical arrangement of components. The presented system includes four stages; keyboard mapping, impact detection, posture recognition and interaction. Initially keyboard layout is mapped into the keys at the keyboard mapping stage. During the impact detection stage, any interaction performed by the user is identified. Posture recognition stage is used to identify the hand postures that are reserved for the mouse interaction and determine whether the action is an interaction of mouse or keyboard. Interaction stage is finally used to perform the identified interactions by sending the appropriate commands to the system.

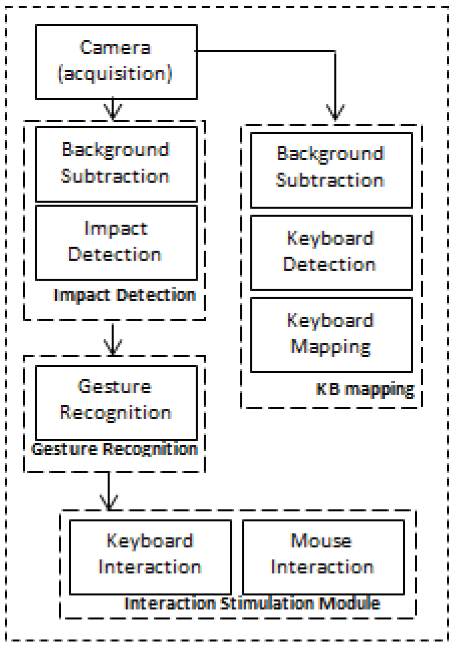

Figure 4: System Architecture

A. Hardware Setup

Hardware setup include a typical low cost web camera, 650nm laser module (which can be found on typical barcode readers) which projects horizontal laser line over the surface and printed QWERTY keyboard layout (optional - because system can also use with augmented reality keyboard). Camera setup should be in a stable position over the display, which is mounted using either a stand or fix into the top of the monitor as illustrated in figure 5. The field of view of the camera should cover complete keyboard layout. Laser beam is also projected parallel to the surface where the beam projects over the surface above 2-3 mm. USB port was used to provide the power requirement for the laser module (3.0 - 4.5 V).

Figure 5: Hardware setup

B. Keyboard Mapping

Thresholding effect is initially applied to the captured image to segment the keyboard layout from the background. Then contour detection is used to identify the contours of the keys and the identified contours were analyzed and mapped to the each key of the keyboard. Users are allowed to switch to the augmented reality keyboard if printed keyboard layout is not detected. In augmented reality keyboard, virtual keyboard layout is augmented to the captured image, enabling users to interact with virtual keyboard.

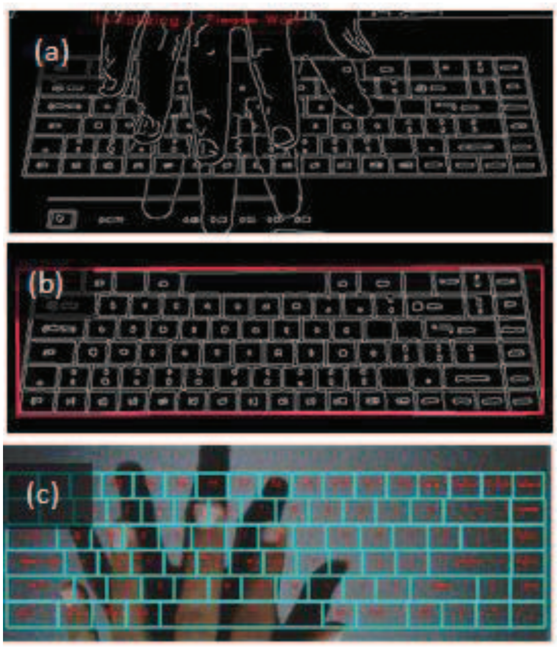

Figure 6: Keyboard detection: (a) Initializing, (b) detected keyboard, (c) augmented reality keyboard

Keyboard contour is assumed to be the biggest square shaped contour detectable in the printed keyboard layout. The positioning of camera is important to detect the keyboard accurately.

After detecting the keyboard layout, all the keys are created as objects and assigned the X and Y coordinates for each key. X, Y coordinates are pixel values of a particular key according to the video resolution. Four coordinate values are used to store the position of each and every key.

$$p_{keyN} = (x_{np}, x'_{np}, y_{np}, y'_{np}) \qquad (3)$$

C. Impact Detection

During this phase activities of the user on the surface were detected using the laser beam. This approach relies on the laser beam impact on fingertips. The colored fingertips are identified using background subtraction and by only selecting color values lay between the color ranges of laser beam.

$$I_{ext} = \begin{cases} 255\ if\ (RGB_{min} \leq RGB_{pix} \leq RGB_{max}) \\ 0\ otherwise \end{cases} \qquad (4)$$

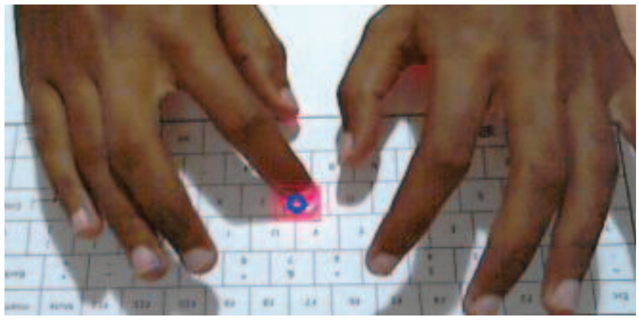

Figure 7: Virtual Keyboard Interactions: sample keystroke

Positive color pixels between the given color range is identified as the laser impact on fingertip and output is given as a set of coordinates which defines the impacts for further processing. These interactions can either be mouse interactions or keystrokes and these are determined according to the postures using Haar cascade classifier.

D. Hand posture detection using Haar-Like features

Set of trained Haar cascades were used for the detection of hand which is in different postures during the interaction. This approach determines whether the interaction is mouse or keyboard interaction. This process only executes when an impact on laser beam ensued and therefore does not affect negatively for the application performance. Three separate Haar classifiers were used to identify three different hand postures. Depending on the number of impacts detected, application selects the appropriate trained cascade to identify the hand posture which is assigned to the particular mouse interaction events. The separations of classifiers for different events significantly reduce the processing time for object detection. This step improves the interaction significantly by avoiding the switching times between the devices. For the training process of each classifier, 500 positives and 1000 negatives were used. Positive samples include positives which were taken under different lighting conditions and different backgrounds to improve the detection rate of the postures under different lighting conditions.

E. Interaction

1) Keyboard Interaction: If the interaction is not identified as a predefined hand posture by Haar classifier, it is identified as a keystroke. The coordinates of the impact is matched against the coordinates of each key object in order to identify which key is pressed. The corresponding key that lies within the coordinates of the touched location will identify and key press event will invoke. When user removes his/her finger, the key up event is fired by the system and release the key.

2) Mouse interactions: If the trained Haar classifiers identify the interaction as one of the three hand postures which is reserved for mouse interaction, the application will switch to the mouse mode and perform the particular event. Single impact on laser stimulates the cursor while two impacts refer to the mouse click events. Index finger is used for the curser while index with thumb is used for left click and index with middle finger used for right click posture. To perform drag operations, same gesture can move through the surface without lifting the fingertips above the surface. Any posture or gesture detected without a laser impact is not considered as a valid mouse interaction. The interaction is valid only if the gesture is valid mouse gesture and detected in the presence of laser impact with the surface.

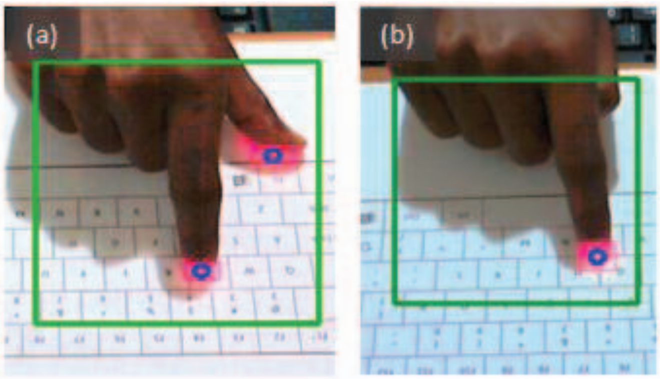

Figure 8: (a) Left click gesture using index and thumb fingers. (b) Pointing gesture using index finger

IV. EXPERIMENTAL RESULTS

In order to evaluate the system, all the components were tested successfully after complete setup arranged according to the given guidelines. The designed application was executed using a workstation with Intel Pentium dual core 2.7 Ghz processor. The keyboard and mouse interactions were able to perform accurately. Few issues were arisen when users pressing multiple keys by intersecting laser beam in same vertical line. Also false impacts were detected when objects with same color of the laser beam is detected on the surface. For experiments, two sets of users were used to operate the system. Since this was a novel way to interact with the computer, most users were not feeling natural and familiar to use the proposing system and that also caused some erroneous interactions during the testing process. Therefore to effectively evaluate the system, two sets of users with (user set 1) and without (user set 2) prior interaction experience were used.

	False detections	**Missed strokes**	**Incorrect detection**	**Typing accuracy**	**Averag e typing speed (wpm)**
User set 1	3.6%	7.3%	5%	84.1%	55.2
User set 2	6.5%	13.2%	9.2%	71.1%	36.4
Average	**5.05%**	**10.25%**	**7.1%**	**77.6%**	**45.8**

Table 1: System usability test summary

Performances of the trained Haar classifiers were also evaluated. A sample of 100 positive images of each posture which were taken in different light conditions was used for this evaluation.

Posture	**Hits**	**Missed**	**False**
Pointer	92	8	17
Right click	88	12	9
Left click	95	5	14

Table 2: Performance of trained classifiers

Test results were at satisfactory level to prove that the application is able to provide accurate interaction by using low cost setup without wearing any color markers.
Demonstration available at: http://youtu.be/41xqdi73gQ4

V. CONCLUSION

A laser based marker-less interactive surface using Haar-like features which enables mouse and keyboard interactions is presented. The proposed system used single typical web camera and a horizontal laser projection module with printed keyboard layout to stimulate all the traditional keyboard and mouse interactions without the need of wearing any markers. Thus the research successfully presents a low cost interactive surface mechanism that provides markerless natural interaction, with the satisfactory results of its evaluation. Few issues were identified in the posture recognition when interacting in poor light conditions. The remaining challenges are to improve the accuracy and to overcome the issues identified with the multiple key presses.

VI. FUTURE WORK

The accuracy and the robustness of the posture recognition can be improved by training the Haar classifier with more positive and negative samples. Also Haar classifier can be trained to detect more postures to assign to other interactions such as zooming, closing, maximizing etc. Infrared laser diode can be also used for the laser projection to avoid the false detection occurred due to the objects of the same color range of the currently used laser module.

VII. ACKNOWLEDGEMENT

Authors wish to thank Department of Computer Science of University of Middlesex, all staff members at International College of Business and Technology and to everyone who supported throughout the study by providing their valuable comments and involvement. Authors also wish to acknowledge Mr: Ranga Abeysooriya, Lecturer, Department of Textile and Clothing Technology, University of Moratuwa, for his guidance and support.

REFERENCES

[1] M. Kölsch, and M. Turk, “Keyboards without keyboards: A survey of virtual keyboards*”, in proceedings of Sensing and Input for Media-centric Systems,* 2002

[2] S. Malik, and J. Laszlo, “Visual touchpad: A two-handed gestural input device, *“in Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04)*, pp: 289-296*,* 2004.

[3] P. Mistry, (2013, July 12), Sixthsense [online], Available: http://www.pranavmistry.com/projects/sixthsense/ .

[4] M. Lee, and W. Woo, “*ARKB:* 3D vision-based Augmented Reality Keyboard,” *in Proceedings of International Conferece on Artificial Reality and Telexisitence,* 2003.

[5] Huan Du *et al*., “A Virtual Keyboard Based on True-3D Optical Ranging,” *in Proceedings of the British Machine Vision Conference*, vol. 1, p. 220 – 229, 2005.

[6] N. Conic, P. Cerseato, and De Natale, F. G. B.,: “Natural Human-Machine Interface using an Interactive Virtual Blackboard”, *In Proceeding of ICIP*, pp.181-184, 2007.

[7] S. Rautaray & A. Agrawal, "Real time multiple hand gesture recognition system for human computer interaction," *J. Intelligent Syst. Applications*, vol. 5, pp. 56–64, 2012.

[8] P. Viola, M. Jones, “Robust real-time object detection,” *Cambridge Research Laboratory Technical Report Series CRL2001/01*, pp. 1-24, 2001.

[9] Q. Chen, N.D. Georganas, E.M. Petriu, “Realtime vision-based hand gesture recognition using haar-like features,” *In Proceedings of. IEEE Instrument and Measurement Technology Conference,* 2007.

[10] P. Viola, M. Jones, “Rapid object detection using a boosted cascade of simple features,” *In Proceeding of CVPR*, pp. 511-518, 2001.