Real Time Webcam based Infrared Tracking for Projection Display System

Aniket Kudale\textsuperscript{a}, Kirti Wanjale\textsuperscript{b}

\textsuperscript{a} ME Computer Engineering Student, VIIT, Department of Computer Engineering, Pune-411028, India
\textsuperscript{b} Associate Professor, VIIT, Department of Computer Engineering, Pune-411028, India

Abstract

In this paper, we propose an interaction based projection display system, which enables an infrared pen touch interaction on flat surface (e.g. walls, whiteboard, tables), with a webcam (with an infrared filter removed) and a projector. The challenge of infrared pen touch detection is to sense the touching and movement information of an infrared pen on the surface just from the 2-dimensional image captured by the webcam. In our method, the content of a computer is projected on the surface and the user interacts with the surface using an infrared pen, the movement of an infrared pen is captured by the webcam. Here, the infrared LED light act as a tracking point which helps in controlling the content projected on surface. The proposed method is performed in three stages: 1) the infrared pen LED image is captured and tracked using webcam 2) the calibration of webcam-projector setup is done by principle of Homography 3) the tracked coordinates or location of Infrared pen LED is mapped to computer cursor. By this way, movement of infrared pen allows user to control computer content projected on the surface.

Index Terms: Human Computer Interaction, Camera-Projector System, User Interface, Infrared Tracking.

© 2016 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science

1. Introduction

Recent advancement in the technology has created great opportunity to improve human life. In Human Computer Interaction (HCI) [6], users interact with computers through keyboard and mouse which requires physical touch to these devices. But, camera based perceptual system provides an alternative way to interact with computer. In this system, generally a webcam is used to track the movement made by user and the tracked movement is then translated to the motion parameters to operate computer. The field of Computer Vision, Human Computer Interface (HCI), and Perceptual User Interface (PUI) has provided a way to ease the interaction with computing systems.

* Corresponding author. Tel.: +91-9850440272
E-mail address: aniket.kudale@hotmail.com
There are various possibilities available for interacting with computing devices, other than our traditional devices mouse and keyboard [14][19]. Today computers have become faster and more reliable, so it is possible to introduce a novel or alternative solution to control computer. The solution can be achieved by making computers to capture some movement made by user or some perceptual motions which can be captured by sensing device such as webcam. Basically, it can be done by tracking down the movement made, by capturing the live video feed from webcam and processing it. The computer the processes the acquired information and interprets it into action or response. This methodology or technology is referred as perceptual user interface technology. The most important part in HCI is tracking or sensing user’s input, most of the time webcam or other camera device is used to track the users movement [7], the movement is then converted to motion parameters which are then translated to computer operations [12]. This method can be illustrated by following framework Fig. 1.

![Basic Framework](image)

This methodology can be applied in a novel way to create useful applications, and introduce a new way of Human Computer Interaction paradigms. Currently, the interactive white boards (IWB) are been widely used in various organizations and educational institutes [3]. Interactive smart boards are defined as touch-sensitive boards that operate through a computer with typical overhead projectors. There are basically four main components used to realize such system; computers, digital projectors, an electronic board, and a software tool that in compatible with hardware setting. The interaction with the system can be done in various ways such as resistive, electromagnetic, capacitive, ultrasonic and certain optical based sensors. Moreover to these methodologies, there are different techniques which acts as signal source, categorized into two as active marker and passive marker. Active markers have better stability while passive markers are unstable for detection. Most of the commercial smart boards are controlled by the use of finger, stylus or other devices. Some use new techniques for interaction such as natural gestures. Natural gesture as captured by natural interaction devices such as Kinect [10], in order to interact with computers. Such methodologies can be adapted to build interactive smart board.
This paper is organized as follows. Section 2 summarizes the related works which includes survey of touch detection methodologies in projection display systems. Section 3 proposes the system along with the description of hardware setup used and system architecture. Section 4 presents the implementation details such as algorithms and techniques used in the system. In section 5, we present our experimentation results followed by conclusion in section 6.

2. Related Works

The first interactive whiteboard was used in 1991, but in recent years such whiteboard have seen widely in high-tech classrooms. Basically, an interactive whiteboard consist a dry erase whiteboard with a projector mounted on wall. With software, the whiteboard becomes computer screen. The projector projects the computer content onto the board surface where user controls the content either with stylus or touch of the hand instead of mouse and keyboard. Whatever operations are done on computer monitor is replicated on the interactive whiteboard. In educational scenario, a teacher can create or teach interactively the lessons such as matching activity where students have to match items, either using finger or stylus. Research findings have proved that students understand and learn better when they are fully engaged and multisensory, hands-on learning is the best way to engage them. Interactive whiteboard can maximize the use of limited computer access. Students can collaborative work together with other individuals contributing at the board, other participants at the computer. In recent years, there has been lots of research going on in developing interactive whiteboard surfaces. Hu, Jun, et al. [1] proposed a novel approach which enables finger touch interaction by detecting touch event on projected surface based on button distortion. This requires a standard camera and a projector, and can be used to emulate interactive whiteboard based on finger touch. Chen Rui, et al, [2] proposed an IR pen-based table top system that uses Wii-remote to track the IR pen movement and thus interact with the projected content. In this system they developed a novel method allowing multiple IR pens to be detected and tracked simultaneously at the same time. Dai, Jingwen, et al. [5] proposed a system that lets any table top surface to which content is projected to be a touch enabled surface, using just a projector and a camera. Touch is detected by using computer vision algorithms on hand/finger such as hand segmentation and tracking. Soares Christope, et al. [6] proposed a simple camera, projector system and an IR pen for interactive whiteboard operation. They used webcam with photographic film attached to it, in order to detect IR pen light. This system however faced accuracy issue as background noise disturbs the functioning and also affects its efficiency. Chien-Yu Lin, et al. [8] discussed the application of Wii-remote based interactive surface in educational settings Michal Lech, et al. [9] developed a system which allows interactive whiteboard to be controlled using hand gesture of user. Presented a hardware and software on virtual whiteboard with special focus on utilizing the fuzzy rule-based gesture recognition concept and kalman filters for the prediction of consecutive hand positions also evaluated their system for efficacy. Shuai Zhang, et al. and Marco Ronchetti, [10][11] proposed interactive whiteboard setup using Kinect gaming system as a sensing device. Kinect is connected to computer which acts as a sensing device to locate the pointer such as hand, IR pen on board on which the content is projected. Problem with such system is the cost of the Kinect device. Massimo Bosetti, et al, [13] presented a result analysis on usage Wii-remote based interactive whiteboard by conducting a survey in educational institutes to understand various concerns and issues regarding usage of Wii-remote for interactive whiteboard. Pranav Mistry, [14] introduced wearable gesture interface. It lets user to project content on any surface and user can interact with it using hand gesture [15]. A marker such as a color tape is wrapped around the fingers and a portable projector camera setup sense the color marker gesture made by fingers [19] to interact with the board. Johnny Chung Lee, [16] [17] proposed the first Wii-remote based interactive whiteboard. The system consist of a projector, a computer, a Wii-remote connected to computer via Bluetooth; this acts as a sensing device, an Infrared pen is used as a pointer; to interact with the projected content on board surface. Andrew Wilson, [20] proposed a compact interactive tabletop projection vision system. It uses camera-projector pair. Introduced touch detection algorithm based on the observation of shadows also used an optical flow algorithm for interacting with onscreen information. Francois Brard, [21] emulated interactive whiteboard...
using camera projector pair. Colored plastic tokens were used to interact with the projected content on board surface. Color detection algorithms are used in the system to detect the position of color tokens on surface of whiteboard. Kudale, et al. [26] reviewed all available such systems.

3. Proposed System

We propose an interaction based projection display system in which projector projects computer’s desktop on board surface where user can interact with the content projected on surface using infrared LED pen Fig. 2. Multiple pens can be also used for collaborative work by multi users.

Fig.2. System Overview

3.1. Hardware Prototype

The hardware prototype consists of projector, a webcam with IR filter removed, an Infrared LED pen and a computer. The projector, is connected to computer, which projects computer content on plane surface (e.g. whiteboard, walls), an infrared LED pen is use to interact with the content projected on plane surface, the webcam, with IR filter removed, captures the movement made by infrared LED pen and computer analyzes the captured image to detect touch interactions. The infrared pen is Infrared LED connected to a battery in series with momentary switch.

The main components of the system are: First, Canon REALiS LE-5W Projector [24]. Its native resolution is 1280x800 (WXGA) with focus adjustment which allows screen sizes from 20 inches up to 240 inches. Second, Logitech C170 webcam [23] that provides 1024 x 768 resolution image recording at 30FPS. The viewing area of the webcam is comparatively larger than that of projector, which ensures that the content projected on the surface is able to be captured by the web camera. Finally, we use laptop (Intel Core i7 @ 2.80GHz CPU, 8.00GB Dual-Channel DDR3 @ 664MHz RAM) to analyze and process captured image.

3.2. Webcam IR Filter Removal

Basic webcam has capability of detecting Infrared LED light source, but the high luminance of the IR light is blocked by the IR filter which is in webcam, this makes infrared LED light source barely detectable. This is
Real Time Webcam based Infrared Tracking for Projection Display System

...done because the CCD or CMOS in web cameras are highly sensitivity to near-infrared light. In our proposed system we are tracking only the infrared LED light using webcam. We have removed the IR filter [25] from the webcam and have replaced it with floppy disk magnetic coating material which is made up of 1.2 µm cobalt doped iron oxide. This helps in minimizing the extra Infrared LED Light source intensity to be captured by webcam also it eliminates the need of background subtraction algorithms while tracking the movement of infrared LED light and makes the detecting and tracking less complex.

![System Architecture](image)

Fig.3. System Architecture
3.3. System Architecture

The system basically consists of five modules: computer vision, motion tracker, Homography, region-of-interest tracker and mouse mapping as shown in Fig. 3.

The computer vision module will detect the infrared blob in the video feed by using image processing and computer vision algorithms. A webcam is used as sensing device which captures the motion made by infrared led pen. Once the infrared blob is detected its center coordinates are determined and the detected coordinate values are passed onto motion tracker modules.

The motion tracker module will track the X, Y coordinates of infrared blob. This is important step for Homography module. Coordinates obtained from this module will be used for calibration in Homography module.

4. Implementation Details

4.1. Infrared LED Pen Detection Algorithm

The first step is to detect Infrared LED light via webcam. Each image captured through webcam is processed in order to find high intensity Infrared LED Light. This is done using image processing algorithms. The Image captured is first converted to Grayscale by this formula

\[(R + G + B) / 3\]  

Each RGB color point contains red R, green G, and blue B color elements. These color points are converted to gray scale using the above equation. After grayscale conversion of image, it is converted into binary (i.e. Black and White) image. In image Binarization process a threshold value is selected, then all pixels with values above this threshold are set white, and all pixels with values less than threshold value are set to black. Here to choose optimal threshold value, adaptive image Binarization technique is used.

Blob detection is a technique used to detect infrared LED light it determines if groups of connecting pixels are similar to each other. This is useful for identifying unique objects in a scene, or counting the number of objects in a scene. To detect an infrared LED light blob, we put threshold to the image by a specific color, white in our case. If there is only one blob in a scene, the middle mass is always located in the center of an object. We used one component at a time algorithm to detect blob in an image.

**ALGORITHM:** Blob Detection based on connected component labeling

1. go through each pixel in the array:
2. if the pixel is a blob color, label it ’1’
   otherwise label it 0
3. go to the next pixel
4. if it is also a blob color and if it is adjacent to blob 1
   label it ’1’
   else label it ’2’ (or more)
5. repeat until all pixels are done

What the algorithm does is that, it labels each blob by a number, counting up for every new blob it encounters. Then to find middle mass, you can just find it for each individual blob.

4.2. Calibration of Webcam and Projector

Calibration is required to map the coordinate system of webcam to the coordinate system of the projected screen. Infrared LED light source pen plays a vital role in calibration process, it is a one-time process. When
the system is started for the first time, an image with four points on it at four corners is displayed. User uses infrared LED pen and presses the on/off switch, system detects the event and triggers a function, and i.e. location of this point is remembered and at the end the average of these points are calculated. This method is repeated for other points as well. These points are in the coordinate system of the webcam.

4.3. Homography Principle

Homography principle is a concept in mathematical science of geometry [22] Fig. 4. A Homography is an invertible transformation from projective plane to a projective plane that maps straight lines to straight lines. This is required for calibration process. This module maps the coordinates of projected screen to computer screen. Thus creating a region, in which only the infrared blob will be detected and tracked.

For transforming webcam coordinates \( x, y \) to projection screen coordinates \( X, Y \), we will need the matrix \( H \), so given by formula

\[
\begin{bmatrix}
X
Y
1
\end{bmatrix}
= \begin{bmatrix}
X_a / m
Y_a / m
m / m
\end{bmatrix}
= \begin{bmatrix}
h_{11} & h_{12} & h_{13}
h_{21} & h_{22} & h_{23}
h_{31} & h_{32} & h_{33}
\end{bmatrix}
\begin{bmatrix}
x
y
1
\end{bmatrix}
\]

(2)

Now the \( H \) matrix can be written in vector form as \( h = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33})^T \) the homogeneous equations for \( n \) points become \( |A \cdot h| = 0 \) with \( A \) the 2n x 9 matrix. The values of \( x, y \) are known, but the \( H \) matrix is unknown. This is computed by the following formula \( |A \cdot h| = 0 \), where the \( A \) matrix is defined as:
Here the vector $\mathbf{h}$ that minimizes the residual $|A \cdot \mathbf{h}| = 0$, subjected to $|\mathbf{h}| = 1$, is given eigenvector of eigenvalue of $A^T A$. This eigenvector can be obtained by singular value decomposing $A$. If suppose $n = 4$, is the null vector of $A$ and the residuals are 0.

$$A_{2n \times 9} = U_{2n \times 2n} S_{2n \times 9} V^T_{9 \times 9}$$

(4)

$$V^T V = I$$

(5)

$$U^T U = I$$

(6)

The solution will be the row of matrix $V$. The no. of row is determined by the count of rows of the $S$ matrix with lowest value, which will be divided into 3 rows by 3 numbers to create a new $H$ matrix.

4.4. Mouse Cursor Control

After detecting the Infrared LED blob from an image, its position/coordinates are computed. The computed coordinate’s values are passed to mouse API, which helps in the movement of cursor on screen. Click event is triggered when the user presses the on/off switch on pen. When no light is detected in region-of-interest, a function is called; this indicates a click operation which will be carried by mouse API of an operating system.

5. Experimental Results

5.1. Performance

This system can track Infrared LED located at distance approximately 5 meters. The resolution for tracking is 1024 x 768 pixels and the display resolution depends upon the projector and computer settings. We tested our system on 1024 x 768 resolution and were able to achieve decent frame rate of 20-25 fps using Logitech C170 Webcam.

5.2. Infrared LED Detection Test

In our initial experiment we also tested webcam’s capability to detect infrared LED light. Blob detection algorithm is implemented to detect Infrared LED light source. We tested this algorithm on three different
webcam settings. First, we tested the algorithm implementation on Logitech C170 webcam with no modifications shown in Fig. 5.

![Fig.5. Infrared Detection Test on Webcam with No Modifications.](image)

As you can see, the user is holding an infrared LED light source, but the current webcam setup (no modifications) is not able to detect the infrared LED source, instead its detecting other moving objects i.e. in this figure the user’s hand, the light falling on user is getting reflected and is similar in intensity to infrared LED light. We are interested in tracking only the infrared LED light source, this can be done on current webcam setting by integrating this approach with background subtraction technique, which removes the background, but it takes more time and computing power and is not accurate.

Then, we tested the same algorithm on slightly different webcam setup, we added a photographic film [6]. This minimizes the extra light which is reflected off the user’s body movement. This setup helps in detecting infrared LED light source more effectively than previous setup. Fig. 6. Shows the output from current setup.

![Fig.6. Infrared Detection Test on Webcam with Photographic Film Attached](image)
As you can see, the user is holding an infrared LED light source, the current webcam setup (with photographic film attached) is able to detect the infrared LED source, but it is also detecting other moving objects i.e. in this figure the user’s hand, this may also require integrating this approach with background subtraction technique, which removes the background, this will be more accurate than previous approach.

Finally, we tested the same algorithm webcam setup with IR filter removed and replacing it with Floppy disk material. This approach allows only infrared LED light source to be detected, as floppy disk material blocks the extra light source and since IR filter is removed, sufficient intensity of infrared LED light source can be detected. Fig. 7. Shows the output from current setup.

As you can see in the above figure, the user is holding an infrared LED light source, the current webcam setup (with IR filter removed) is able to detect the infrared LED source, and not detecting other moving objects i.e. in this figure the user’s hand is not visible, Hence, there is no need to integrate this approach with background subtraction technique, which removes the background, this will be highly accurate and more efficient than previous approaches. Table I. summarizes these tests and its outcome.

Table 1. Testing Blob Detection Algorithm on Different Webcam Setup.

<table>
<thead>
<tr>
<th>Webcam Setup / Settings</th>
<th>Infrared LED Light Detected</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webcam with no modifications</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>Webcam with photographic film attached [6]</td>
<td>Yes</td>
<td>Moderate</td>
</tr>
<tr>
<td>Webcam with IR filter removed and replaced with floppy disk material</td>
<td>Yes</td>
<td>High</td>
</tr>
</tbody>
</table>

5.3. Robustness

The light emitted from the Infrared LED, must be consistent, in our tests, the IR LED did not remain consistent, it varied in size and intensity as user may hold the pen at different angle, which may cause instability and may not get detected. In order to achieve good detection accuracy rate, webcam must not get obstructed by any other object while detection, in other words IR LED must be in line of sight of webcam.
5.4. Latency in Detection

We observed that if the frame rate of the webcam dropped to 10 fps, detection rate is low. This can be improved by using high resolution webcam. But, there will be cost trade-off.

5.5. Compatibility

We tested our system on different operating environment. The change in the environment does not affect the working of the system. We successfully tested our system on Windows 7 and Linux Ubuntu 13.04.

5.6. Comparison with similar systems

We compared our prototype with other systems based on following criteria.

- Cost: This evaluates the cost of all equipment used, on scale low (less than 10$), medium (20$) and High (>20$). Computer and Projector cost is not included in this.
- Compatibility: Compatibility with Operating system.
- Equipment: List of components required for setup.
- User Friendliness: Ease of using the system on scale easy, moderate, and difficult.

Table 2. Comparison with Similar System.

<table>
<thead>
<tr>
<th>System</th>
<th>Cost</th>
<th>Compatibility</th>
<th>Equipment</th>
<th>User Friendliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our Prototype</td>
<td>Low</td>
<td>Windows, Linux</td>
<td>Webcam with IR Filter removed, IR Pen</td>
<td>Easy</td>
</tr>
<tr>
<td>Johny Lee Whiteboard</td>
<td>Medium</td>
<td>Windows</td>
<td>Wiimote, Bluetooth adapter, IR Pen</td>
<td>Moderate</td>
</tr>
<tr>
<td>Kinect Whiteboard</td>
<td>High</td>
<td>Windows</td>
<td>Kinect</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

The above table shows that our prototype comparatively low in cost, compatible with windows and linux platform and also easy to use.

6. Conclusion

We proposed a system, which tracks infrared LED light source using a webcam with IR filter removed for projection display system. System architecture specifies how the system can be used by the user to control the content projected by projector on surface using an Infrared LED pen. We have developed our system using OpenCV and C++. We also described how a webcam can be modified to detect infrared LED light source only which can be used effectively in our system. This approach reduces the need using high cost infrared detecting camera and also, requires less processing since only infrared LED is detected and tracked. We evaluated our system on parameters robustness, infrared detection on three different settings, latency, compatibility and comparison with other available system. This approach can be used in various infrared tracking applications such as low cost interactive whiteboard, touchscreens in educational institutes or corporate companies.

Acknowledgement
I would like to express my gratitude to my research project guide Prof. K. H. Wanjale for valuable suggestions and remarks. I would also like to thank our department of computer engineering for giving me this opportunity.

References


Canon Projector - http://www.usa.canon.com/cusa/consumer/products/projectors/portable_projectors/_le_5w_bk


Author’s Profiles

Aniket E. Kudale is a post graduate student, pursuing masters degree in computer engineering in Vishwakarma Institute of Information Technology (VIIT), affiliated to Savitribai Phule Pune University, Pune. His research area includes Computer Vision, Web Technologies, Human Computer Interaction and Open Source Technologies.

Mrs. Kirti Wanjale is currently working as an Associate Professor in the Department of Computer Science in Vishwakarma Institute of Information Technology (VIIT), affiliated to Savitribai Phule Pune University, Pune. She has 14 years of teaching experience. She is currently pursuing PhD. She has presented papers in many national and international conferences and published articles in many international journals. Her research area is image processing.

How to cite this paper: Aniket Kudale, Kirti Wanjale,"Real Time Webcam based Infrared Tracking for Projection Display System", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.2, No.4, pp.41-53, 2016.DOI: 10.5815/ijmsc.2016.04.05