IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 1, 2018, accepted February 20, 2018, date of publication March 6, 2018, date of current version March 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2812756

Turn Any Display Into a Touch Screen Using Infrared Optical Technique

JI-REN LIANG[1], (Member, IEEE), KUAN-CHIN LIN[1], HSIANG-CHIEH CHEN[2], (Member, IEEE), AND WEN-JUNE WANG[1], (Fellow, IEEE)

[1]Department of Electrical Engineering, National Central University, Taoyuan 32001, Taiwan
[2]Department of Electrical Engineering, National United University, Miaoli 36063, Taiwan

Corresponding author: Hsiang-Chieh Chen (chc@nuu.edu.tw)

ABSTRACT Touch screens have become increasingly popular and widely used in most human–computer interfaces. As screen sizes have become larger for commercial use, the cost-performance ratio has been an important issue for both producers and consumers. In this paper, we present a new low-cost touch display method using an optical technique. Our approach is suitable for any regular screen without size limitation and represents a cost-effective solution. The proposed system employs a single webcam and an infrared laser line projector and incorporates image processing and computer vision methods to achieve the desired touch function. The experimental results of the proposed method provide a novel, real-time, and low-cost system to transform a conventional non-touch display into a touch-sensitive screen. The infrared optical technique can implement touch interactivity on desktop screens or large projection screens. Additionally, users can set up the system easily.

INDEX TERMS Optical technique, touch detection, infrared touch system, touch-sensitive display.

I. INTRODUCTION

Since mobile phones and tablet personal computers (PCs) are widely used in daily life, users find that a touch screen is an essential interface for controlling or instructing a computer. Most recent mobile phones are touch screen devices, which enable the users to interact directly with the display. By enabling intuitive interaction with mobile phones and tablet PCs, the touch screen shortens the feedback time without requiring a mouse and keyboard. One obvious example of the benefits of using a touch screen is that it makes learning easier and more convenient. Typically, there are three types of touch screen techniques in the market, including the resistive, capacitive, and surface wave types. The resistive and capacitive types are widely used in small-size displays such as mobile phones and tablet PCs. However, they are rarely used in large-size displays since the cost of buying a brand new touch screen is high. Consequently, a regular display is often modified to become a touch-sensitive screen due to the budget issue.

Various approaches in touch-sensitive implementation have been represented in many studies. A webcam-based touch screen system is based on a simple and generic finger detection technique. Only a single webcam is used to achieve this goal [1]. A novel interactive projection system enables bare-finger touch interaction on regular planar surfaces, with only one ordinary camera and one projector. A pico-projector can be used to significantly increase the limited screen size of mobile devices [2]. An algorithm of detecting multiple touch points on an imaging surface was implemented using the Java programming language and achieved detection of multiple fingertip blobs [3]. A webcam-based touch screen was proposed in which the main tools are image processing techniques. Two cameras are employed for acquiring a sequence of image frames; the touched object is recognized and its location is identified [4]. A projector and a camera were used to let any tabletop surface to which the projection is illuminated become a touch-sensitive computer screen [5]. A novel 3D pointing system enables interaction with large displays using a single webcam [6]. Another approach to finger detection employs a type of camera-based interactive board. The finger-finding is confined within a stripe that is the projection of the edge of the board on the image plane with respect to a camera instead of using global search. The region where a finger intersects with the stripe is first detected and segmented from the background. A region-growing algorithm is then applied to the region to extract the whole finger [7].

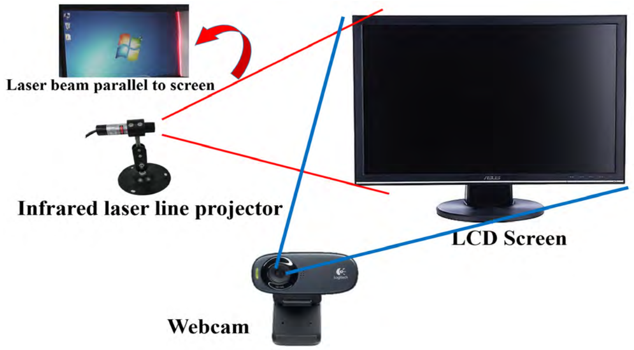


FIGURE 1. Infrared optical system.

FIGURE 2. Detection area on fingertip.

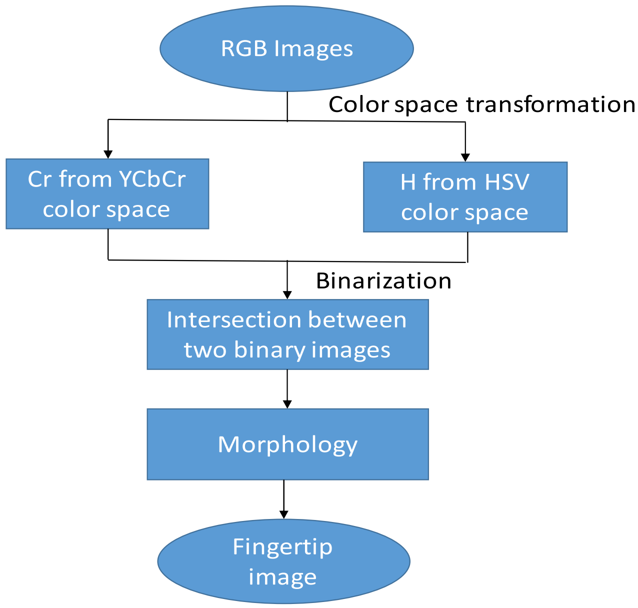


FIGURE 3. Flowchart of image pre-processing.

An interactive visual display system called Magic Crystal Ball was presented. With Magic Crystal Ball, users can manipulate the display with touch and hover interaction [8]. The detected hand and finger can be used to implement a non-contact mouse. This technology can be further used to control home devices such as curtains and televisions. Skin color is used to segment the hand region from background and a counter is extracted from the segmented hand. Analysis of the counter provides the location of the finger-tip in the hand [9]. An intelligent desk system allows a user to perform freehand drawing on a desk or similar surface with gestures. The system requires one camera and no touch sensors. The key underlying technique is a vision-based method that distinguishes drawing gestures and transitional gestures in real time, avoiding the need for artificial gestures to mark the beginning and end of a drawing stroke [10].

This paper proposes a novel method of utilizing a camera and an infrared laser line projector to enable the touch-screen functionality on any display. The geometric position of a display to camera and projector are not necessarily fixed, and it can be localized with an automatic localization system. Additionally, this system can transform a regular display into a touch-sensitive screen without any expensive facilities, which is especially advantageous for large displays. In brief, the proposed technology allows any normal screen to have touch features, and it has the advantages of low price, simple installation, and ease of use.

II. SYASTEM CONSTRUCTURE OVERVIEW

The prototype of the proposed infrared optical system is illustrated in Fig. 1. The system consists of three parts: (1) A single webcam with HD video capturing (up to 720p under 30 frames per second) that is used to capture the motion of the finger with a laser spot on it. (2) An infrared laser line projector with wavelength 650nm. The laser beam must be close to and parallel with the regular screen to ensure that the laser line can form a spot (detection target) on the fingertip. (3) A desktop computer and a 23-inch LCD screen for the experiments. The system was configured for an unfixed geometric position that is discussed in the next section. When the fingertip touches the screen surface, it shows the detection target with the infrared light beam illuminating on it, as shown in Fig. 2. The webcam is located on the same side of the projector to capture the screen edges and fingertip images. Note that the position of the projector and webcam influence the touch-sensitive performance.

A. IMAGE PRE-PROCESSING

The procedure of image pre-processing is plotted in Fig. 3. Because the laser beam emits to the finger and infrared light appears on the fingertip, the pixel luminance plays an important part to discriminate between the fingertip and the background. In our experimental design, the image of fingertip is easier to localize and preprocess than whole hand. Accordingly, the original RGB image captured from the webcam is first converted to the images represented in YCbCr and HSV color formats. Because of the need to control the effects of environmental light, we choose V component from HSV color space; because the laser beam is red light, we choose Cr

FIGURE 4. Binary image of the fingertip.

FIGURE 5. Red, blue, and white screen colors.

component from YCbCr color space. Then, we binarize those two color images and next to choose the intersection part between the two binary images. In our experiments, we take the intersection part, and then set a threshold value equals 10 to turn the color images into binary images. The fingertip appears only as white points in the image. By employing morphological erosion and dilation, the fingertip is reserved as our detection area. Fig. 4 shows the detection result of Fig. 2.

B. AUTOMATIC POSITIONING

Without the restriction of webcam position, we need to detect the screen edges for forming the touching area. An automatic positioning method is proposed to localize the screen edges. Since the YCbCr color space can take the chrominance of red and blue colors, we manually display the screen in three different colors (red, blue, and white), as shown in Fig. 5. First, the red and white color images are converted to YCbCr color spaces, and we take the difference of each Cr part of the images. Then we obtain the red color screen image. We do the similar way for the blue and white color images. Fig. 6 shows the intersection of detection results of red and blue screen images (left), and the result after noise removal by morphological erosion and dilation. The final step is to perform edge and corner detection to localize the screen edges and corners as shown in Fig. 7.

C. COORDINATE TRANSFORMATION

Although the screen edges and corners in the image are well localized, the coordinates in the captured frame cannot directly represent the real coordinates related to the screen. There exists a transformation from image to real world coor-

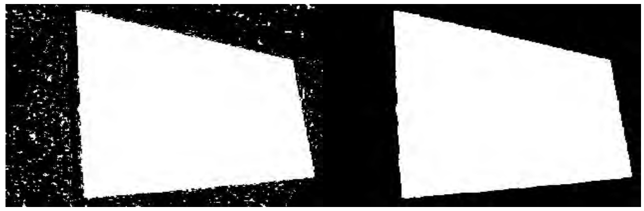


FIGURE 6. Intersection of two binary image (left), and after morphology (right).

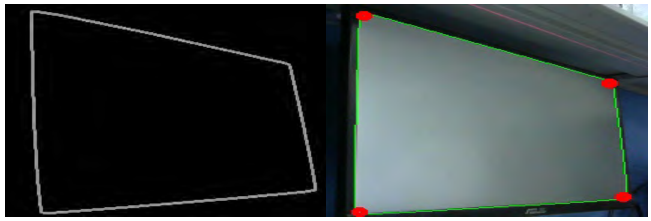


FIGURE 7. Edges and corners localization on screen.

dinate system. A transform matrix can be utilized to achieve this coordinate mapping. In this work, we use a 3 × 3 transform matrix as a projection matrix and state that

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \tag{1}$$

where (x_1, y_1) is the coordinate of the fingertip in the captured image, and (x_2, y_2) is the coordinate after coordinate mapping. Then, we can use the coordinates of the four corners in the captured image to map the corners of the screen. We assume the four coordinates of the captured image as (x_{11}, y_{11}), (x_{12}, y_{12}), (x_{13}, y_{13}), (x_{14}, y_{14}), and each coordinate maps to the corners of the screen as (x_{21}, y_{21}), (x_{22}, y_{22}), (x_{23}, y_{23}), (x_{24}, y_{24}).

We can use (1) to derive from the first set (x_{11}, y_{11}) and (x_{21}, y_{21}), then we can derive

$$\begin{cases} ax_{11} + by_{11} + c = x_{21}\text{.........(a)} \\ dx_{11} + ey_{11} + f = y_{21}\text{........(b)} \\ gx_{11} + hy_{11} + i = 1\text{............(c)} \end{cases} \tag{2}$$

Then we use (2c) times x_{21} and y_{21} separately, and we can obtain the functions as

$$\begin{aligned} gx_{11}x_{21} + hy_{11}x_{21} + ix_{21} &= x_{21}\text{.......(a)} \\ gx_{11}y_{21} + hy_{11}y_{21} + iy_{21} &= y_{21}\text{......(b)} \end{aligned} \tag{3}$$

By combining (2a) with (3a), and (2b) with (3b), we can derive the functions as

$$\begin{aligned} ax_{11} + by_{11} + c &= gx_{11}x_{21} + hy_{11}x_{21} + ix_{21}\text{.........(a)} \\ dx_{11} + ey_{11} + f &= gx_{11}y_{21} + hy_{11}y_{21} + iy_{21}\text{........(b)} \end{aligned} \tag{4}$$

and turn (4) to (5) as

$$\begin{aligned} ax_{11} + by_{11} + c - gx_{11}x_{21} - hy_{11}x_{21} &= ix_{21} \\ dx_{11} + ey_{11} + f - gx_{11}y_{21} - hy_{11}y_{21} &= iy_{21} \end{aligned} \tag{5}$$

The rest of the three sets repeat the same procedures, and we can also derive the functions as

$$\begin{aligned} ax_{12} + by_{12} + c - gx_{12}x_{22} - hy_{12}x_{22} &= ix_{22} \\ dx_{12} + ey_{12} + f - gx_{12}y_{22} - hy_{12}y_{22} &= iy_{22} \end{aligned} \quad (6)$$

$$\begin{aligned} ax_{13} + by_{13} + c - gx_{13}x_{23} - hy_{13}x_{23} &= ix_{23} \\ dx_{13} + ey_{13} + f - gx_{13}y_{23} - hy_{13}y_{23} &= iy_{23} \end{aligned} \quad (7)$$

$$\begin{aligned} ax_{14} + by_{14} + c - gx_{14}x_{24} - hy_{14}x_{24} &= ix_{24} \\ dx_{14} + ey_{14} + f - gx_{14}y_{24} - hy_{14}y_{24} &= iy_{24} \end{aligned} \quad (8)$$

The functions (5)-(8) can be represented in the matrix form as

$$\begin{bmatrix} x_{11} & y_{11} & 1 & 0 & 0 & 0 & -x_{11}x_{21} & -y_{11}x_{21} \\ 0 & 0 & 0 & x_{11} & y_{11} & 1 & -x_{11}x_{21} & -y_{11}y_{21} \\ x_{12} & y_{12} & 1 & 0 & 0 & 0 & -x_{12}x_{22} & -y_{12}x_{22} \\ 0 & 0 & 0 & x_{12} & y_{12} & 1 & -x_{12}x_{22} & -y_{12}y_{22} \\ x_{13} & y_{13} & 1 & 0 & 0 & 0 & -x_{13}x_{23} & -y_{13}x_{23} \\ 0 & 0 & 0 & x_{13} & y_{13} & 1 & -x_{13}x_{23} & -y_{13}y_{23} \\ x_{14} & y_{14} & 1 & 0 & 0 & 0 & -x_{14}x_{24} & -y_{14}x_{23} \\ 0 & 0 & 0 & x_{14} & y_{14} & 1 & -x_{14}x_{24} & -y_{14}y_{24} \end{bmatrix} \times \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} x_{21} \\ y_{21} \\ x_{22} \\ y_{22} \\ x_{23} \\ y_{23} \\ x_{24} \\ y_{24} \end{bmatrix} \quad (9)$$

Hence, variables from a to h are then solved by considering (9) as a nonhomogeneous system. Finally, the projection matrix is derived as

$$T = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (10)$$

D. MULTI-FINGERTIPS IDENTIFICATION

When we need to perform a zoom-in or zoom-out motion, the first step is to identify how many fingertips touch the screen. Assume there are two fingers touched the screen as shown in Fig. 8, then after image pre-processing, we can get a binary image as shown in Fig. 9, and the next step is to find out all the white points where the positions are. By checking all the white points in the binary image, we can find out the maximum and minimum white points on x and y coordinates, as shown in Fig. 10). If the distance between these two points is greater than a certain threshold, and it represents there are two fingertips touch the screen at the same time. Here we should notice that the two fingertips cannot point on the same horizontal line, once a fingertip is hide by the other fingertip, it is impossible to form two laser spots. Calculate the distance between the maximum and minimum white points on x and y coordinates. When the distance is getting longer, it represents the fingertips start zoom in motion. Otherwise, it represents the fingertips start the zoom out motion.

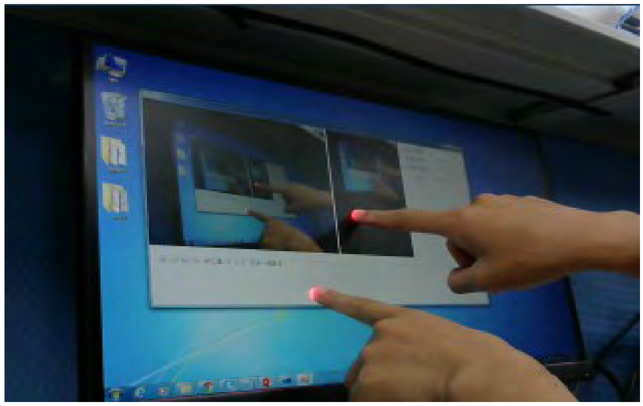


FIGURE 8. Original image captured by webcam.

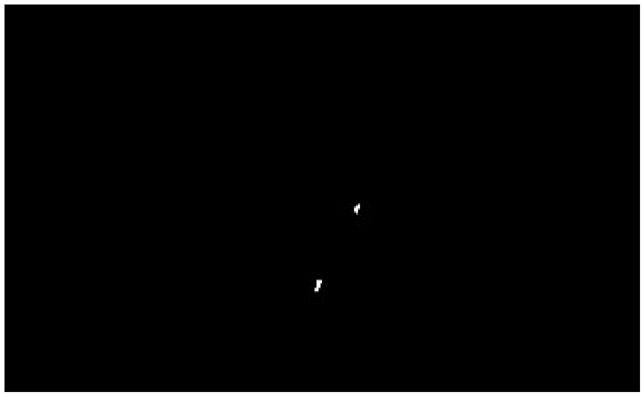


FIGURE 9. Binary image shows two clusters of white points.

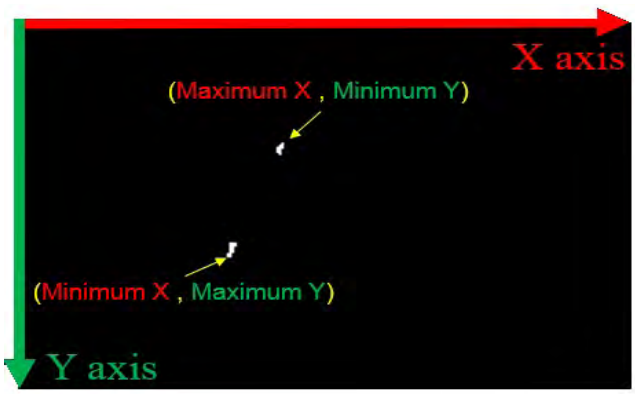
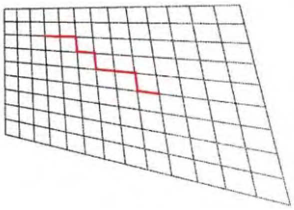


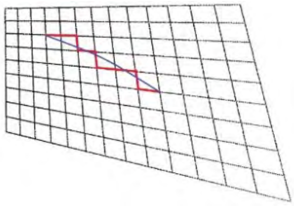
FIGURE 10. The maximum and minimum white points on x and y coordinates.

E. HAND-WRITTEN LINE INTERPOLATION

The resolution used to detect a fingertip image which was captured from webcam is 320 × 240 pixels owing to real-time consideration, and is then mapped to the screen with high resolution such as 1920 × 1080 pixels. The transformation between two different resolutions is not one-to-one. Fig. 11(a) depicts the zigzagging effect while a straight line in low-resolution image is mapped into a high-resolution image. To solve this problem, our work selected a bilinear



(a)



(b)

FIGURE 11. A straight line mapped into the high-resolution image (a) without interpolation, and (b) with interpolation.

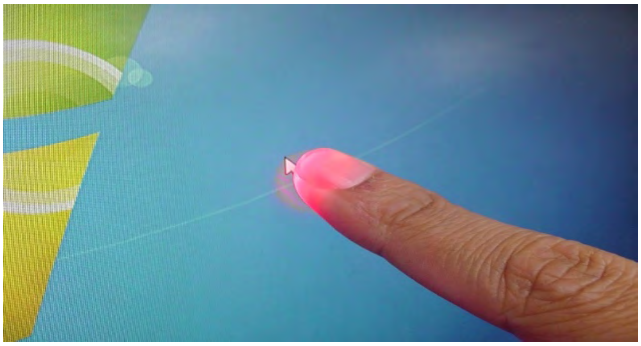


FIGURE 12. Mouse cursor moves with finger.

interpolation method to smoothen the line. The interpolated result is shown as the blue line in Fig. 11(b).

III. EXPERIMENTAL RESULTS

The proposed system has been illustrated in Fig. 1. In this work, we implemented several touch functionalities that are frequently used in touch-sensitive devices. Additionally, the performance on hand writing is compared with the Zhang's method [1] which only employed a single camera to achieve the touch capability.

A. SYSTEM INITIALIZATION

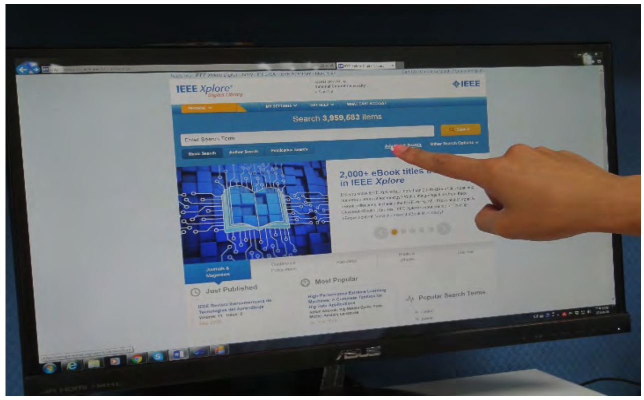
When the proposed system starts, it automatically initializes by the following procedure. First, the screen displays red, blue, and white colors sequentially on the full screen. Then, the positioning is instantaneously completed and the mouse cursor will move by following the finger-tip, as shown in Fig. 12.

B. CLICK FUNCTION TEST

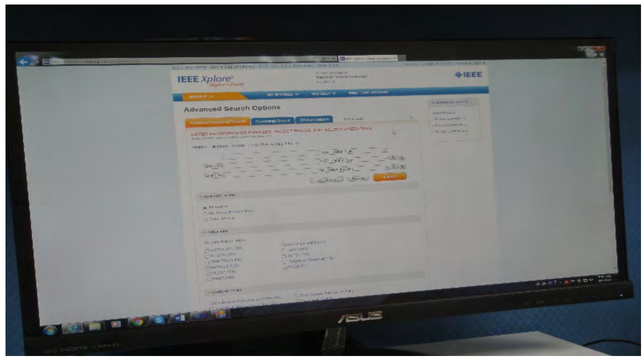
We choose the IEEE Xplore homepage to test the click function, as shown in Fig. 13. When the finger touches the screen for a period of time, then the right-click action of the mouse is triggered, as shown in Fig. 14.

C. ZOOM-IN AND ZOOM-OUT TEST

The proposed system supports the zoom-in and zoom-out gestures. We took Windows Photo Viewer and the news webpage as test samples, as shown in Figs. 15-17. Our gesture performed very well in zoom-in and zoom-out tests.

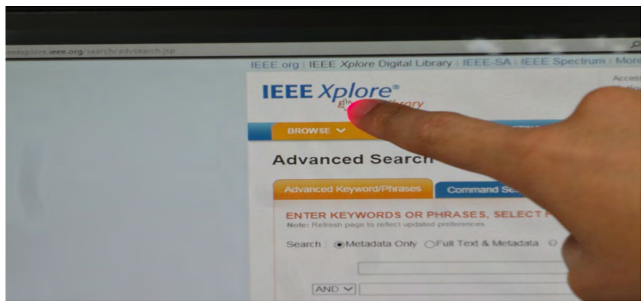


(a)

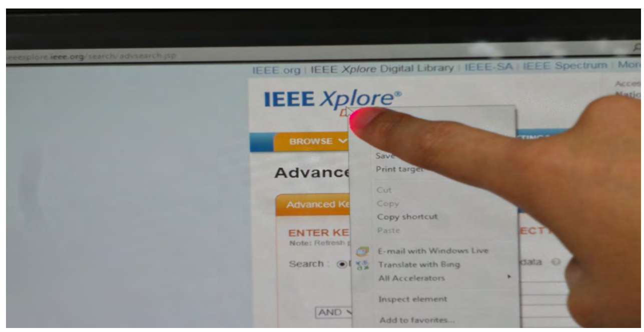


(b)

FIGURE 13. Click function test: (a) Touch a webpage link, and (b) webpage is opened.



(a)



(b)

FIGURE 14. Right click test: (a) Touched the screen for a period of time, and (b) right-click is triggered.

D. PRINT SCREEN FUNCTION TEST

We also defined a gesture for the print screen function (PrtSc). When two fingers are placed on the screen

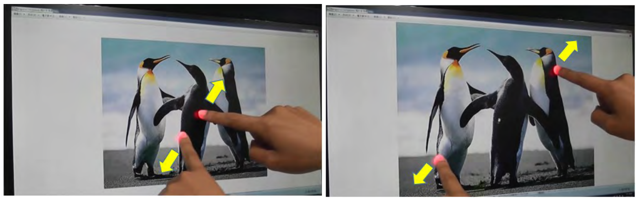


FIGURE 15. Zoom-in test on Windows Photo Viewer.

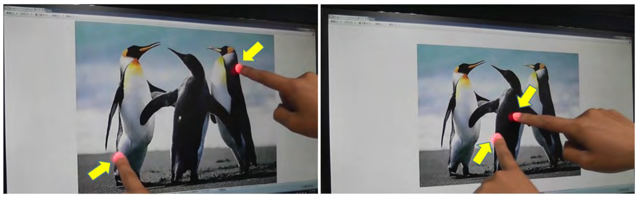


FIGURE 16. Zoom-out test on Windows Photo Viewer.

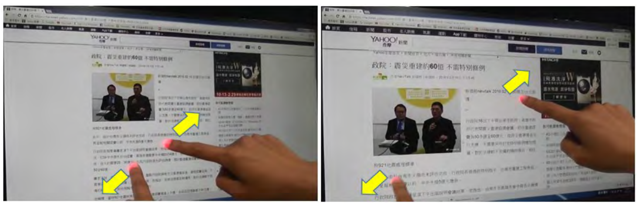


FIGURE 17. Zoon-in and zoom-out text on webpage.

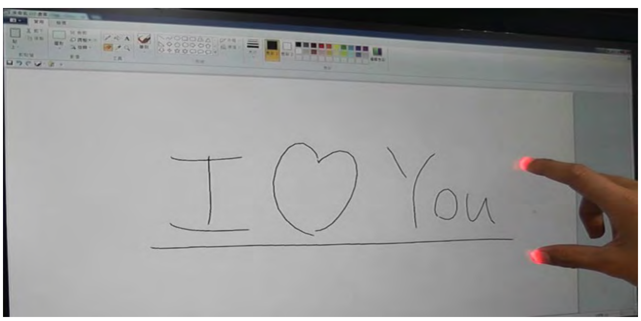


FIGURE 18. Print screen function test by two fingers on the screen for 2s.

for 2s, then the print screen function is triggered, as shown in Fig. 18.

E. CLICK GAME TEST

In addition to normal touch operations, we used the online click rhythm game ''OSU'' to examine the performance on touch sensitivity. The test result shows the touch effect performs quite well, as shown in Fig. 19.

F. DIFFERENT SCREEN SIZES TEST

The above experiments were all tested on a 23-inch display with 1920 × 1080 resolution. In order to verify the proposed system can also work well on different screen sizes, we chose a 15-inch laptop screen with 1366 × 768 resolution to be the test model and performed the same tests as described previously. The experimental results show the proposed system is not affected by the screen size and resolution, as shown in Fig. 20.

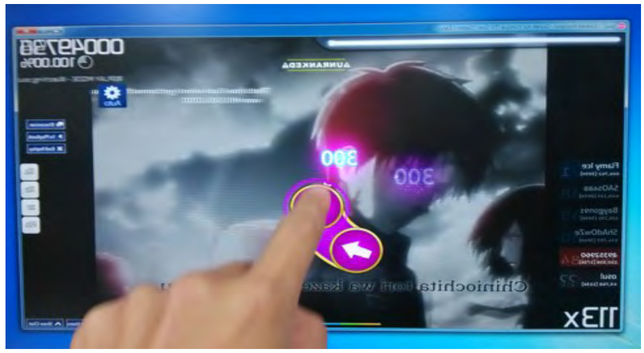


FIGURE 19. OSU online click game touch performance test.

FIGURE 20. Laptop screen size test.

TABLE 1. Angles and distance.

ANGLE BETWEEN PROJECTOR AND SCREEN (Θ)	DISTANCE (D) FROM WEBCAM TO THE SCREEN (M)
$20^\circ \sim 70^\circ$	< 1

G. WEBCAM LOCATION TEST

The proposed system does not restrict the location of the webcam, we still need to avoid anything but fingers block the line of infrared laser or vision of webcam. Every time the system is turned on, it needs to restart the positioning. In our experiment, whether the webcam can capture the whole screen image is the key point to positioning. Table 1 shows the best angles θ between the projector and screen are 20° to 70°, and the best distance d from the webcam to screen is less than 1m, as shown in Fig. 21.

H. EXPERIMENTAL RESULT EVALUATION

Both Zhang's method [1] and the proposed system employ camera-based approaches to transform a regular display into a touch screen. Zhang's method only used a single camera, while our method combined a camera and an infrared laser line projector. An evident advantage is the touch area because of the restriction of a single webcam's field of view. Fig. 22 demonstrates the comparison with Zhang's method and ours. Notably, our touch area covers the full screen.

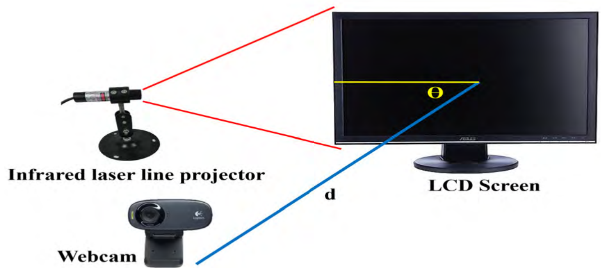


FIGURE 21. Geometric position between webcam, infrared laser line projector, and screen.

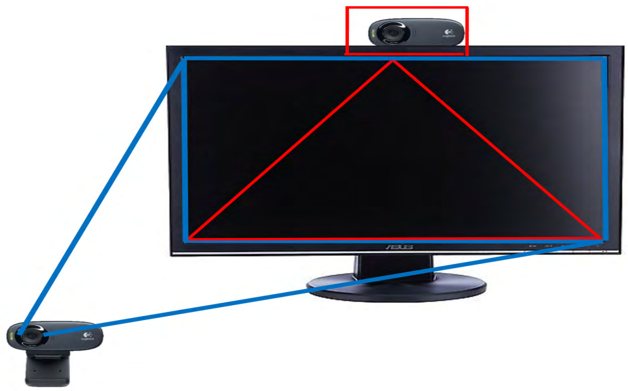


FIGURE 22. Touch area comparison with Zhang's method [1] (Red lines) and our method (Blue lines).

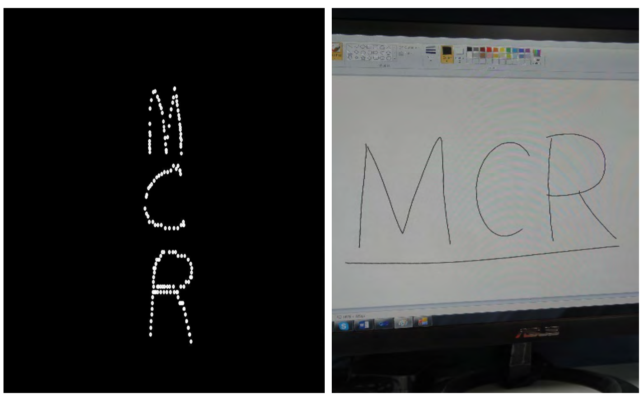


FIGURE 23. Writing comparion between Zhang's method (left) and ours (right).

Furthermore, the writing results were also evaluated in our experiments. Trying to write three alphabets on the screen as shown in Fig. 23, the result from Zhang's method appeared broken segments in each word. Our method showed the superiority to writing because the hand-written interpolation was used.

IV. CONCLUSION AND FUTURE WORKS

The experimental results show that the proposed infrared optical technique has a superior touch-sensitive performance, and it can provide a better user experience. Our approach is suitable for any regular screens without size limitation and turns a regular LCD screen into a touchable display. The system has the benefits of low cost and innovation. In the future, we would like to improve the system stabilization and increase the multi-touch points from two to ten. Additionally, because the ambient light affects the image processing performance, we attempt to develop automatic adjustment scheme of the image parameters to make the system become more robust.

REFERENCES

[1] L. Zhang, J. Saboune, and A. El Saddik, "Transforming a regular screen into a touch screen using a single webcam," *J. Display Technol.*, vol. 10, no. 8, pp. 647–659, Aug. 2014.

[2] J. Hu, G. Li, X. Xie, Z. Lv, and Z. Wang, "Bare-fingers touch detection by the button's distortion in a projector-camera system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 4, pp. 566–575, Apr. 2014.

[3] P. C. Ravoor, S. R. Rupanagudi, and B. S. Ranjani, "Detection of multiple points of contact on an imaging touch-screen," in *Proc. Int. Conf. Commun., Inf. Comput. Technol.*, Mumbai, India, Oct. 2012, pp. 1–6.

[4] H. M. Phuong, N. T. Dzung, and T. M. Hoang, "Implementation of webcam-based touchscreen," in *Proc. 3rd Int. Conf. Commun. Electron.*, Nha Trang, Vietnam, Aug. 2010, pp. 201–206.

[5] J. Dai and C.-K. Chung, "Touchscreen everywhere: On transferring a normal planar surface to a touch-sensitive display," *IEEE Trans. Cybern.*, vol. 44, no. 8, pp. 1383–1396, Aug. 2014.

[6] K. Cheng and M. Takatsuka, "Initial evaluation of a bare-hand interaction technique for large displays using a webcam," in *Proc. 1st ACMSIGCHI Symp. Eng. Interact. Comput. Syst.*, New York, NY, USA, Jul. 2009, pp. 291–296.

[7] Y. Zhou and G. Morrison, "A real-time algorithm for finger detection in a camera based finger-friendly interactive board system," in *Proc. 5th Int. Conf. Comput. Vis. Syst.*, Bielefeld, Germany, Mar. 2007, pp. 21–24.

[8] L.-W. Chan *et al.*, "Gesture-based interaction for a magic crystal ball," in *Proc. ACM Symp. Virtual Reality Softw. Technol.*, New York, NY, USA, Nov. 2007, pp. 157–164.

[9] S. K. Kang, M. Y. Nam, and P. K. Rhee, "Color based hand and finger detection technology for user interaction," in *Proc. Int. Conf. Converg. Hybrid Inf. Technol.*, Aug. 2008, pp. 229–236.

[10] Z. Mo, J. P. Lewis, and U. Neumann, "SmartCanvas: A gesture-driven intelligent drawing desk system," in *Proc. ACM Int. Conf. Intell. User Interfaces*, Jan. 2005, pp. 239–243.

JI-REN LIANG received the B.S. degree in mechanical engineering from the Chung-Cheng Institute of Technology, Taoyuan, Taiwan, in 1999, and the M.S. degree in electrical engineering from the Chung-Cheng Institute of Technology, National Defense University, in 2005. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electrical Engineering, National Central University. His current research interests include image processing, target tracing, and pattern recognition.

KUAN-CHIN LIN received the B.S. degree in electrical engineering from National Taiwan Ocean University, Keelung, Taiwan, in 2014, and the M.S. degree in electrical engineering from National Central University, Taoyuan, Taiwan, in 2016. He is currently with Taiwan Semiconductor Manufacturing Company with a focus on semiconductor manufacturing.

HSIANG-CHIEH CHEN received the B.S. degree in engineering science from National Cheng Kung University, Tainan, Taiwan, in 2002, and the M.S. and Ph.D. degrees in electrical engineering from National Central University, Taoyuan, Taiwan, in 2005 and 2009, respectively. From 2010 to 2014, he was with the Industrial Technology Research Institute, where he was involved in intelligent video analytics and energy saving. From 2014 to 2015, he was with Altek Semiconductor with a focus on stereo vision research and development. From 2015 to 2016, he served as a Technical Leader of computer vision with Coretronic Corporation and was in charge of innovative technology development. Since 2016, he has been with the faculty of the Department of Electrical Engineering, National United University, Miaoli, Taiwan, where he is currently an Assistant Professor. His research interests include image processing, computer vision, artificial intelligence, and robotics.

WEN-JUNE WANG (F'08) received the B.S. degree from the Department of Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 1980, the M.S. degree from the Department of Electrical Engineering, Tatung University, Taipei, Taiwan, in 1984, and the Ph.D. degree from the Institute of Electronics, National Chiao Tung University, in 1987. He was the Dean of the College of Electrical Engineering and Computer Science, National Central University, Taoyuan, Taiwan. He was also a Chair Professor and the Dean of the Research and Development Office, National Taipei University of Technology, Taipei, from 2007 to 2009. From 2005 to 2007, he was the Dean of the College of Science and Technology, National Chi-Nan University, Nantou, Taiwan. He is currently a Chair Professor with the Department of Electrical Engineering. He has authored or co-authored over 160 refereed journal papers and 153 conference papers in fuzzy systems and theorems, robust and nonlinear control in large-scale systems, and neural networks. His most significant contributions are the design of fuzzy systems and the development of robotics. His current research interests include robot control, neural networks, and pattern recognition.