

Bare Finger 3D Air-Touch System Using an Embedded Optical Sensor Array for Mobile Displays

Guo-Zhen Wang, Yi-Pai Huang, Tian-Sheuan Chang, and Tsu-Han Chen

Abstract—Generally, 3-D interactive systems are used in systems with large displays and are not suitable for portable devices. To “touch” a floating 3-D image displayed from a mobile device, we proposed a camera-free, 3D, interactive, virtual-touch system for bare fingers. By embedding optical sensors into the display pixels and adding angular scanning illuminators to the edge of the display, a flat panel can sense the images reflected off a bare finger. The three-axis (x, y, z) information of the fingertip can be determined by analyzing these reflected images. The proposed 3-D virtual-touch system successfully functions with a 4-in mobile display.

Index Terms—Three-dimensional (3D) virtual touch, embedded optical sensors.



Fig. 1. Concept of playing a virtual 3D STACKO game on a 3D mobile display using air-touch technology.

I. INTRODUCTION

TOUCH systems are interface between users and machines. Most current mobile displays are limited to a 2-axis (x and y) positioning for objects [1]. These axes are insufficient for 3D displays [2]–[5] or other 3D interactive applications. Many methods and devices, such as the WorldViz PPT (precision position tracking) [6], the Wii, etc., have been used to achieve 3D interactivity. However, none of these technologies are applicable to mobile display applications. A 3D mobile display with 3D virtual touch functionality can provide a more intuitive and friendly interface, as shown in Fig. 1.

Current 3D interactive systems primarily fall into machine-based and camera-based groups. Machine-based systems [7] require wearing additional devices to detect motion, and can provide feedback vibrations to the users. However, these systems are considered inconvenient because they require additional heavy devices to be worn.

For camera-based systems [8]–[12], 3-axis (x, y, z) positions can be calculated using various methods. For instance, the popular Kinect [13] can detect relative 3-axis (x, y, z) positions using two infrared (IR) cameras with a corresponding IR light source. However, this camera-based system requires high resolution images to calculate the 3-axis (x, y, z) positions, and the resolution is proportional to the size of camera sensor; therefore,

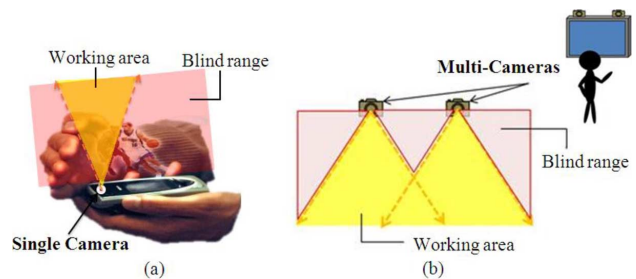


Fig. 2. The limited field of view for both (a) single-cameras and (b) multi-cameras causes a blind range within the continuous interaction space.

the form factor impedes such systems from being integrated into a portable device. The Leap 3D motion control system was recently unveiled by a San Francisco startup called “Leap Motion” and can provide a 3D hands-free motion controller. However, the Leap Motion technology still requires an additional device and is thus not suited to mobile devices.

Camera-based systems are usually limited by their field of view (Fig. 2) and form factor, which prevents them from detecting objects in close proximity to the display and being integrated into portable devices such as smart phones and tablet.

Another example is the BiDi screen that was proposed by M. Hirsch in 2009 [14]. The BiDi screen uses a sensor layer a small distance from the normal LCD display. When this sensor layer views an object through the optical mask, information about the distance the object is in front of the screen can be captured and decoded. This technology is generally known as depth from focus [15]. Because a BiDi screen uses an optical tiled-MURA mask to capture information about the object, it relies on the environmental brightness. Without sufficient lighting, the device cannot capture a clear image. Furthermore, the mask decreases the detected brightness, which increases the required ambient conditions.

For application in real portable devices, an embedded optical sensor based approach [16]–[18] was proposed that can be integrated into the display pixels to provide light and maintain a

Manuscript received April 01, 2013; revised July 31, 2013; accepted August 01, 2013. Date of publication August 07, 2013; date of current version December 27, 2013. This work was supported by the National Science Council of Taiwan through Academic Projects NSC 101-2221-E-009-120-MY3.

G.-Z. Wang and T.-S. Chang are with the Department of Electronic Engineering, National Chiao Tung University, Hsinchu 300, Taiwan.

Y.-P. Huang is with the Photonics Department, Display Institute, Hsinchu 300, Taiwan.

T.-H. Chen is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850 USA.

Color versions of one or more of the figures are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JDT.2013.2277567

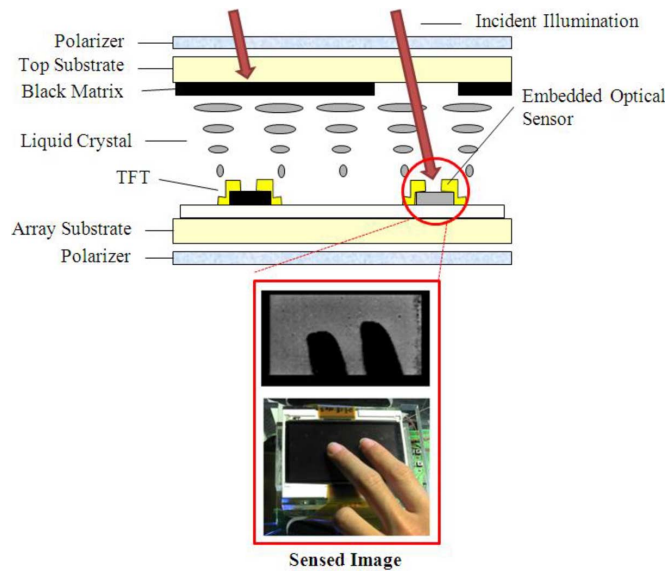


Fig. 3. Schematic structure of an optical sensor embedded onto a TFT substrate and the sensed image.

thin form. As shown in Fig. 3, the embedded optical sensors are not blocked by the black matrix, thus a photo current will occur when the sensor receives light. The embedded optical sensor was first proposed for 2D touch applications [19] and has been recently extended for 3D interactions [20]. However, the device still required a light pen for previous applications [21], [22]. To achieve a bare finger 3D air-touch device, we proposed an embedded optical sensor system with angular scanning illuminators added to the display edges, as shown in Fig. 4. The proposed system allows for near-distance 3D air-touch functionality above the entire surface of the display. Additionally, it can be very compact and independent of ambient conditions. Therefore, floating 3D images can be touched and interacted with.

II. CONSTRUCTION AND METHOD

The proposed bare finger 3D virtual touch system is constructed from a traditional display, embedded optical sensors, infrared (IR) backlight, and angular scanning illuminators. To calculate the 2-axis (x and y) positions of a fingertip, the IR backlight passes through the panel and is reflected by the fingertip as shown in Fig. 4(a). To calculate the depth (z) of the fingertip, the IR angular scanning devices are placed on opposite sides of the panel as shown in Fig. 4(b).

Unlike previous image capturing techniques, which are strongly dependent on ambient illumination, the proposed system uses embedded IR sensors and light sources. The de-background image processing of each detected frame allows the device to operate under various ambient conditions.

The flow chart of the proposed algorithm is shown in Fig. 5. At first, the raw data is retrieved from embedded optical sensors. For reducing noises from environment and system, noise suppression which includes de-noise and de-background will be adopted.

Second, intensity of captured image from IR backlight needs to be accumulated. When the accumulated value is larger than touch threshold value, which is determined by the calibration

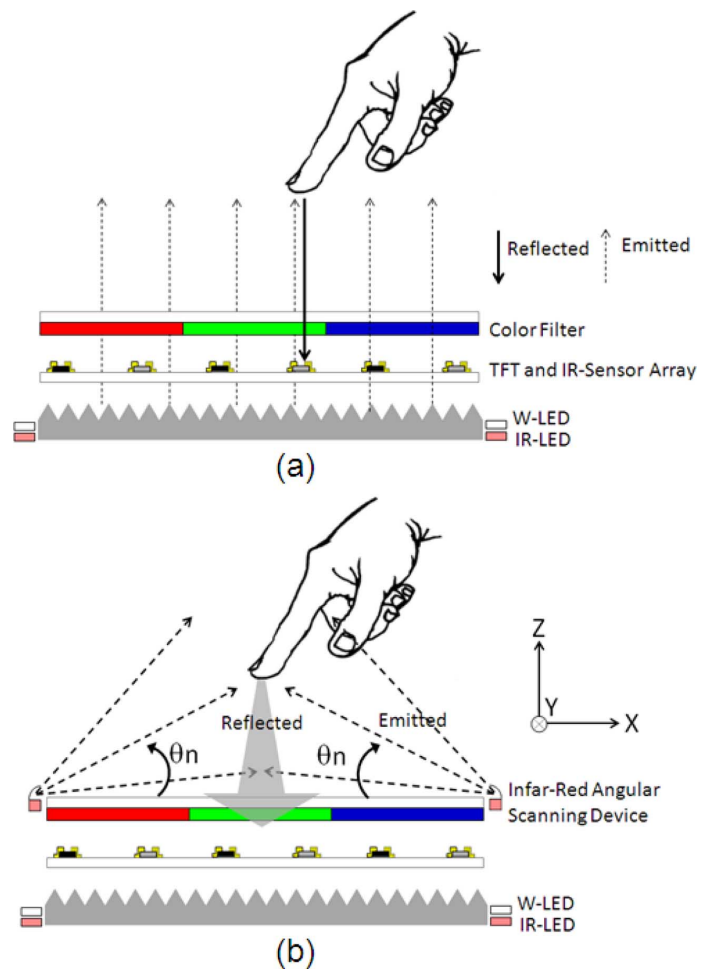


Fig. 4. Cross section of the bare finger 3D virtual touch system in (a) IR backlight illumination mode (for x and y determination) and (b) IR angular scanning mode (for z determination).

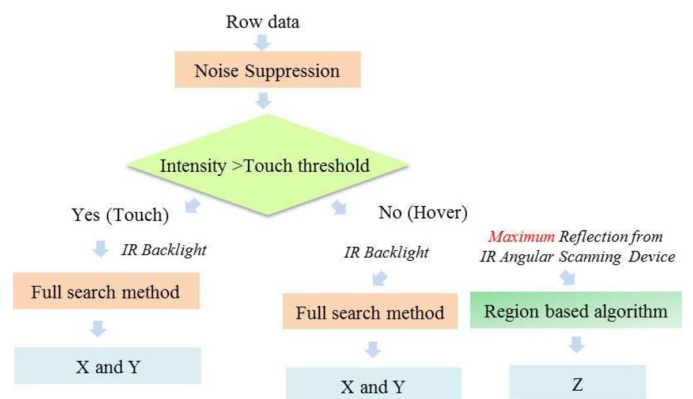


Fig. 5. Flow chart of the bare finger 3D air-touch algorithm.

data, the object will be sensed as touching on the panel. Then the “full search method” can be used to determine the 2-axis (x and y) position of a touch point. Finally, if less than the threshold, the object is hover of the screen, thus the proposed “region based algorithm” will be induced to calculate the depth (z) information. The details of “region based algorithm” will be described in the following.

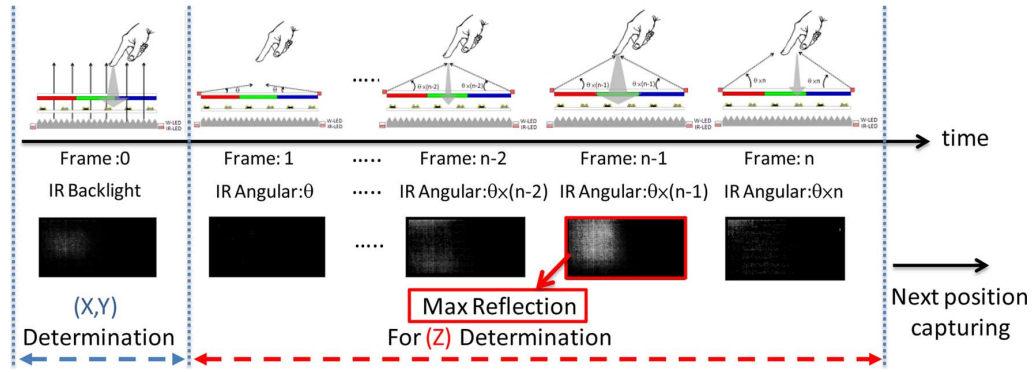


Fig. 6. Timing diagram for the bare finger 3D virtual touch system using an embedded optical sensor and two angular scanning illuminators.

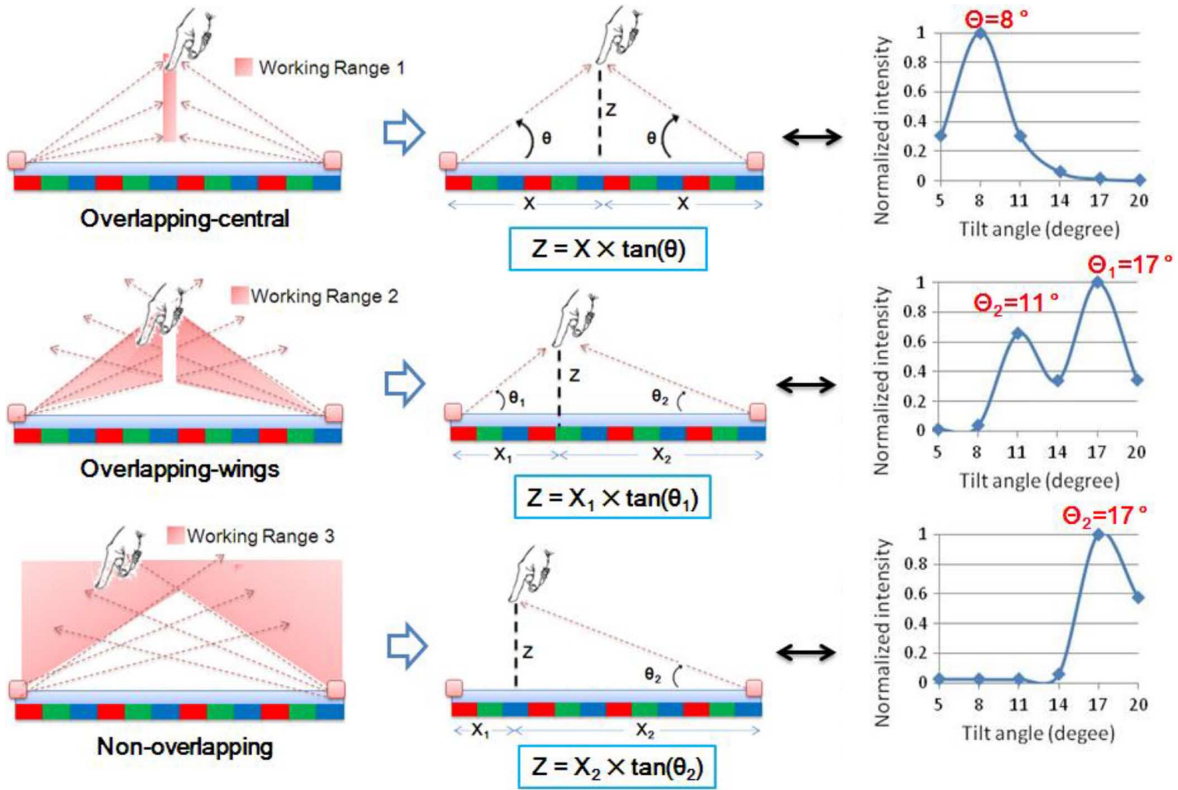


Fig. 7. Concept for the region based algorithms in the different working regions.

Next, we present the timing diagram (Fig. 6) for the proposed bare finger 3D virtual touch system to clearly illustrate the working process. Initially, the IR backlight and IR angular scanning illuminators will need to be synchronized with the embedded optical sensors. In the first sensing frame (Frame 0), the IR backlight passes through the panel twice and the optical sensors capture any reflected light to determine the 2D position (x and y) of the fingertip. Next, the IR angular scanning illuminators on the both sides of the panel will sequentially emit light at different tilt angles. These sensing frames (Frame 1 to Frame n) correlate to a tilt angle from θ to $n \times \theta$. Finally, analyzing the accumulated intensity of each frame provides the scanning angle with the maximum reflectance and thus the location of the fingertip. From the 2-axis (x and y) position and scanning angle (θ), the depth (z) of the fingertip can be calculated. Finally, the 3D virtual touch coordinates (x , y , z) can be obtained. In the

proposed system, the algorithm for calculating the 3-axis (x , y , z) position of the fingertip is less complex than for image processing.

To calculate the 2-axis (x and y) position of a fingertip, the full search method is used to analyze the reflected IR backlight image. The full search method [9] uses a filter covered by the image to total the intensity and determine the position with the highest accumulated intensity, which is the 2-axis (x and y) position of the object.

In the following steps, the images from the IR angular scanning can be processed using a region-based algorithm to obtain the depth (z) of the fingertip as shown in Fig. 6. This region-based algorithm exploits the proposed design concept to divide the hover region into three working ranges, 1 (overlapping-central), 2 (overlapping-wings), and 3 (non-overlapping) as illustrated in Fig. 7. In the working range, 1

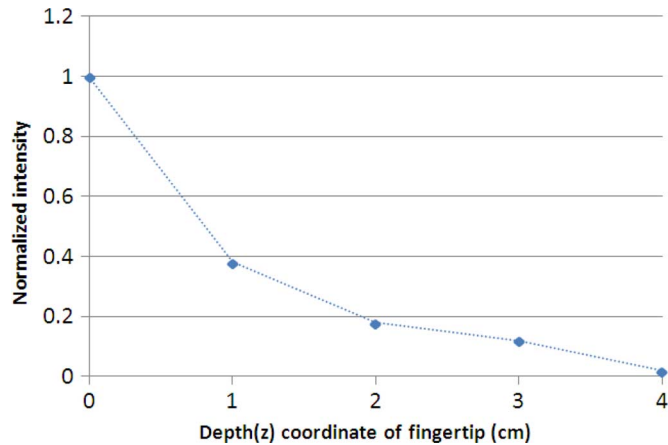


Fig. 8. Measurement results for the reflected intensity of a fingertip at different depths.

(overlapping-central), the fingertip is centered over the display and reflects the two side illuminators at the same tilt angle (θ). Therefore, the intensity of the tilt angle curve only has a single peak for the scanning angle (θ) with the maximum reflectance from the fingertip. When the fingertip is in working range 2 (overlapping-wings), it reflects the two side illuminators at different tilt angles (θ_1 and θ_2). Therefore, the intensity vs. tilt angle curve has two peaks. With working range 3 (non-overlapping), the curve has only a single peak because the fingertip is beyond the scan angle of the closer illuminator. Finally, with both the scanning angle (θ) and 2-axis (x and y) position of the fingertip, the depth (z) information can be calculated using a simple trigonometric function ($z = x \times \tan \theta$).

The proposed approach is suitable for touching and interacting with floated 3D images from a mobile display. Furthermore, the proposed approach has the advantages of providing near-field support for full screen detection, simple algorithms for real-time calculations, and built-in illuminators for independence from ambient conditions.

III. RESULTS

A prototype 3D air-touch system was built using a 4-inch LCD equipped with IR sensors integrated onto the TFT substrate. The sensor resolution was 68×120 , which is one fourth the image resolution of the display. For implementing the proposed concept, the high collimated light and scanning mirror was used to simulate the sequential lighting source. Furthermore, the sequential lighting sources were synchronized with optical sensors.

To verify the accuracy of the 3-axis (x, y, z) positioning of a fingertip above the 4-inch mobile display, a test of x/y and depth (z) coordinates ranging from 1–4 cm and 0–3 cm, respectively. The depth range was limited by the low sensitivity of the current embedded optical sensors. As shown in Fig. 8, the reflected IR intensity decreased to almost zero when approaching 4 cm in height. The experiment results shown in Figs. 9 and 10 indicate the detected coordinates of X/Y and the depth (z) matched the real value. The average error was less than 0.2 cm (x/y) and 0.3 cm (z). Furthermore, the maximum error in 3-axis (x, y, z) direction will, of course, increase with increasing depth (z) as

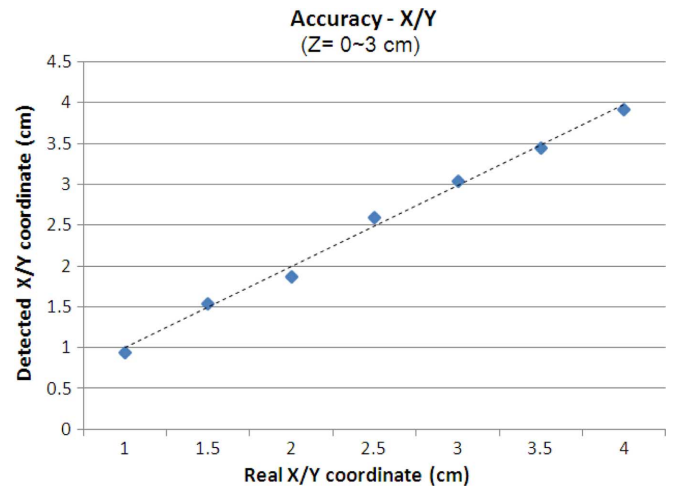


Fig. 9. Accuracy of the x/y coordinates at different depths (z) for the object ($z = 0 \sim 3$ cm).

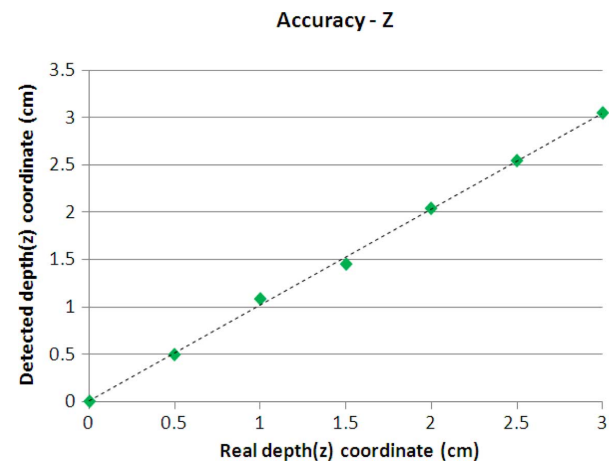


Fig. 10. Accuracy of the depth (z) of the object at different x/y coordinates.

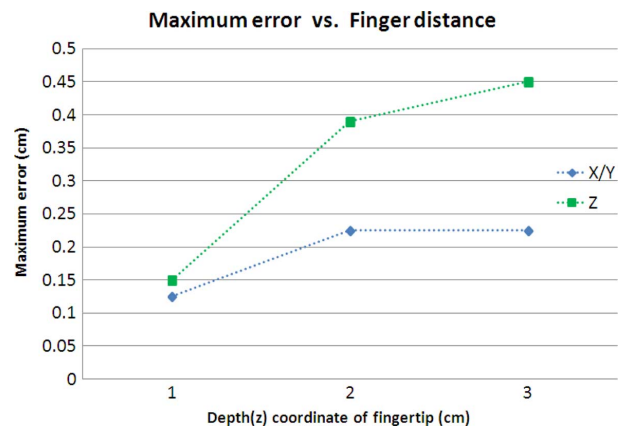


Fig. 11. Maximum error for the x/y and z coordinates at different depths (z) for the object ($z = 1 \sim 3$ cm).

shown in Fig. 11. This trend was because of the scattering effect of the fingertip and the increased scanning gap. However, in our experiment, the maximum error of the depth (z) was 0.45 cm, which can satisfy the Win 7 standard (0.5 cm increment) for 2D touch.

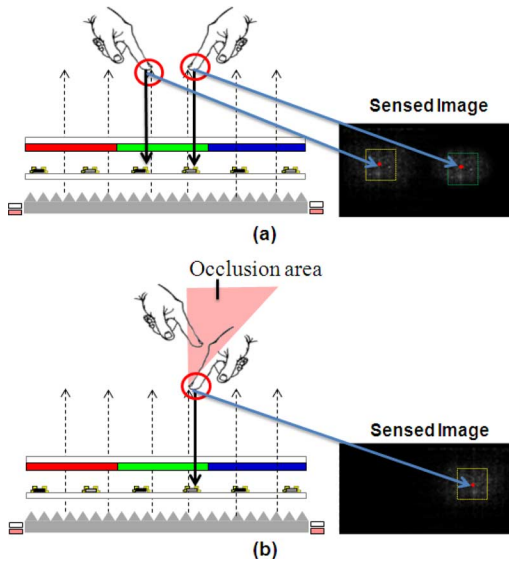


Fig. 12. Comparison of 3D multi-touch in the IR backlight illuminating mode (for the x, y determination) under (a) normal and (b) failing circumstances.

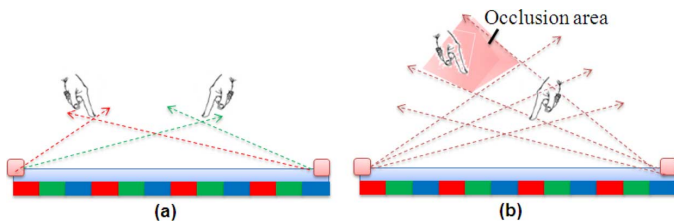


Fig. 13. Comparison of a 3D multi-touch system in the IR angular scanning mode (for the z determination) under (a) normal and (b) failing circumstances.

IV. DISCUSSIONS

To summarize, the approach described above focuses on a 3D virtual single-touch interface. However, including multi-touch functionality could enable more applications for the 3D user interface. Such an interface could be workable with the proposed system; however most multi-touch applications will fail due to the occlusion effect. Fig. 12 shows how the xy -axis of a fingertip in an occluded area cannot be distinguished. Furthermore, if the occlusion area is in the overlapping-wings, the blocked fingertip cannot reflect IR light as shown in Fig. 13, which causes a failure in the multi-touch detection of the depth (z). The interpolation and motion vector methods may be used to overcome the above-mentioned occlusion issues and achieve 3D virtual multi-touch functionality.

V. CONCLUSION

In conclusion, we have presented a camera-free 3D interactive system for providing bare finger air-touch functionality to mobile displays. By embedding optical sensors into the pixels of a display and adding angular scanning illuminators to its edge, a flat panel can sense images reflected from fingertips under most ambient conditions. The sensed images are then used to calculate the 3-axis (x, y, z) position of the fingertip without requiring complex image processing. Finally, the proposed system has been demonstrated using a 4-inch mobile 3D display with a working depth range (z) of up to 3 cm; this range can be further

increased by improving the sensor sensitivity and scanning resolution. However, for touching and interacting with floated 3D images from a mobile display, our approach has the advantages of providing near-field support for full screen detection, simple algorithms for real-time calculations, and built-in illuminators for independence from ambient conditions.

ACKNOWLEDGMENT

The authors would like to express their appreciation to AU Optronics Corporation, Taiwan, for useful discussions in this work.

REFERENCES

- [1] A. D. Wilson, "Playanywhere: A compact interactive tabletop projection-vision system," in *Proc. 18th Annu. ACM Symp. User Interface Software Technol.*, 2005, pp. 83–92.
- [2] J. Y. Son, B. Javidi, S. Yano, and K. H. Choi, "Recent developments in 3-D imaging technologies," *J. Display Technol.*, vol. 6, no. 10, pp. 304–403, Oct. 2010.
- [3] Y. Takaki, "Multi-view 3D display employing a flat-panel display with slanted pixel arrangement," *J. Soc. Inf. Display*, vol. 18, no. 7, pp. 476–482, 2010.
- [4] Y. P. Huang, C. W. Chen, and Y. C. Huang, "Superzone Fresnel liquid crystal lens for temporal scanning auto-stereoscopic display," *J. Display Technol.*, vol. 8, no. 11, pp. 650–655, Nov. 2012.
- [5] Y. P. Huang, L. Y. Liao, and C. W. Chen, "2D/3D switchable autostereoscopic display with multi-electrically driven liquid crystal (MeD-LC) lenses," *J. Soc. Inf. Display*, vol. 18, no. 9, pp. 642–646, 2010.
- [6] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," *Comput. Graph. Appl.*, vol. 22, no. 6, pp. 24–38, 2002.
- [7] S. Feiner, B. MacIntyre, T. Hollerer, and A. Webster, "A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment," in *Proc. 1st IEEE Int. Symp. Wearable Comput.*, 1997, pp. 74–81.
- [8] A. D. Wilson, "TouchLight: An imaging touch screen and display for gesture-based interaction," in *Proc. 6th Int. Conf. Multimodal Interfaces*, 2004, pp. 69–76.
- [9] T. Schlömer, B. Poppinga, N. Henze, and S. Bol, "Gesture recognition with a Wii controller," in *Proc. 2nd Int. Conf. Tangible and Embedded Interaction*, 2008, pp. 11–14.
- [10] T. Schou and H. J. Gardner, "A Wii remote, a game engine, five sensor bars and a virtual reality theatre," in *Proc. 19th Australasian Conf. Comput.-Human Interaction: Entertaining User Interfaces*, 2007, pp. 231–234.
- [11] D. Valkov, F. Steinicke, G. Bruder, and K. Hinrichs, "2D touching of 3D stereoscopic objects," in *Proc. SIGCHI Conf. Human Factors Computing Syst.*, 2011, pp. 1353–1362.
- [12] G. D. Morrison, "A camera-based input device for large interactive displays," *IEEE Comput. Graph. Appl.*, vol. 25, no. 4, pp. 52–57, Jul.-Aug. 2005.
- [13] Lange, B. Flynn, and A. A. Rizzo, "Initial usability assessment of off-the-shelf video game consoles for clinical game-based motor rehabilitation," *Phys. Ther. Rev.*, vol. 14, pp. 355–363, 2009.
- [14] M. Hirsch, D. Lanman, H. Holtzman, and R. Raskar, "BiDi screen: A thin, depth-sensing LCD for 3D interaction using light fields," in *Proc. ACM SIGGRAPH Asia*, 2009, pp. 62–65.
- [15] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graphics*, vol. 26, no. 3, 2007.
- [16] A. Brown, D. Montgomery, J. L. Castagner, H. Kato, and Y. Kanbayashi, "31.3: A system LCD with integrated 3-dimensional input device," in *SID Symp. Dig. Tech. Papers*, 2010, vol. 41, pp. 453–456.
- [17] T. Nakamura, H. Hayashi, M. Yoshida, N. Tada, M. Ishikawa, T. Motai, H. Nakamura, and T. Nishibe, "Incorporation of input function into displays using LTPS TFT technology," *J. Soc. Inf. Display*, vol. 14, pp. 363–369, 2006.
- [18] A. Abileah, W. Boer, T. Larsson, T. Baker, S. Robinson, R. Siegel, N. Fickenscher, B. Leback, T. Griffin, and P. Green, "59.3: Integrated optical touch panel in a 14.1" AMLCD," in *SID Symp. Dig. Tech. Papers*, 2004, vol. 35, pp. 1544–1547.

- [19] W. D. Boer, A. Abileah, P. Green, T. Larsson, S. Robinson, and T. Nguyen, "56.3: Active matrix LCD with integrated optical touch screen," in *SID Symp. Dig. Tech. Papers*, 2003, vol. 34, pp. 1494–1497.
- [20] H. Y. Tung, G. Z. Wang, and Y. P. Huang, "Multi-user and multi-touch system for 3D-interactive display," in *SID Symp. Dig. Tech. Papers*, 2011, vol. 42, pp. 1834–1837.
- [21] G. Z. Wang, M. C. Ma, S. Y. Tung, Y. P. Huang, H. W. Tseng, J. C. Lo, and C. H. Kuo, "A virtual touched 3D interactive display with embedded optical sensor array for 5-axis (x, y, z, θ, φ) detection," in *SID Symp. Dig. Tech. Papers*, 2011, vol. 42, pp. 737–740.
- [22] Y. P. Huang, G. Z. Wang, T. S. Chang, and T. H. Chen, "Three-dimensional virtual touch display system for multi-user applications," *J. Display Technol.*, vol. 9, no. 11, pp. 921–928, Nov. 2013.



Guo-Zhen Wang received an M.S. degree in display institute at the National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 2008 and is currently working toward a Ph.D. at the Department of Electronics Engineering, National Chiao Tung University (NCTU). His current research is to develop 3D interaction systems and focusses on image processing and computer architecture technologies.



Yi-Pai Huang received his B.S. degree from National Cheng Kung University in 1999 and earned a Ph.D. in electro-optical engineering from the National Chiao Tung University in 2004. He is currently a full-time associate professor in the department of photonics & display institute, National Chiao Tung University, Taiwan. He has also been a visiting associate professor at Cornell University from 2011 to 2012. Additionally, he is the chairman of the SID Taipei Chapter, and Chair of the APV program sub-committee, SID. His expertise includes

3D displays and interactive technologies, display optics and color science, and micro-optics. In these fields, he has so far published over 40 Journal and 100 International conference papers and has been granted more than 50 patents. In addition, he has thrice received the SID's distinguished paper award (2001, 2004, 2009).



Tian-Sheuanm Chang (S'93–M'06–SM'07) received B.S., M.S., and Ph.D. degrees in electronic engineering from National Chiao-Tung University (NCTU), Hsinchu, Taiwan, in 1993, 1995, and 1999, respectively. From 2000 to 2004, he was a Deputy Manager with the Global Unichip Corporation, Hsinchu, Taiwan. In 2004, he joined the Department of Electronics Engineering, NCTU, where he is currently a Professor. In 2009, he was a visiting scholar at IMEC, Belgium. His current research interests include system-on-a-chip design, VLSI signal processing, and computer architecture. Dr. Chang received the Excellent Young Electrical Engineer from the Chinese Institute of Electrical Engineering in 2007, and the Outstanding Young Scholar from Taiwan IC Design Society in 2010. He has been actively involved in many international conferences as either part of an organizing committee or technical program committee member. He is current an Editorial Board Member of the IEEE Transactions of Circuits and Systems for Video Technology.



Tian-Sheuan Chen received his B.S. from the National Taiwan University and his M.S. and Ph.D. from Caltech, all in electrical engineering. After working for Bell Labs for several years, he joined the ECE faculty at Carnegie Mellon University in 1997. There, in addition to his research and teaching responsibilities, he served as the associate department head of ECE and the co-director of the Industrial Technology Research Institute (ITRI) Laboratory, a collaborative research program with ITRI in Taiwan. In January 2009, Chen was chosen to become the director of the School of Electrical and Computer Engineering at Cornell University after an extensive national search.