Three-Dimensional Virtual Touch Display System for Multi-User Applications

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

***Abstract*—By embedding optical sensors onto a TFT substrate in a display, a flat panel can sense images projected by infrared (IR) light. Light pens which can project Multi-T-mark including out-mark and in-mark of IR images were utilized to calculate the 3-axis information (x, y, and z) of the tip of the light-pen and achieve accurate user identification accordingly. Therefore, a 3-dimensional virtual touch display for 3-axis information with multi-user/multi-touch system can be achieved.**

***Index Terms*—Embedded optical sensors, multi-marks, three-dimensional (3D) virtual touch.**

I. Introduction

IN THE BEGINNING of 21st century, the wide adoption of smart phones have accelerated a user interface transformation and paved the way to multi-touch technologies [1]. Meanwhile, three-dimensional (3D) technologies [2]–[4] have dramatically changed the user experience. 3D displays [5]–[7] provide additional depth information which generates a more realistic and exciting sensation to users [2]–[8]. Furthermore, augmented reality (AR) technologies which can enhance a user's perception and interaction with the real world are more and more popular [9], [10]. Now most of the AR technologies usually use small marked cards to produce virtual objects [11]. However, the AR technologies cannot provide the depth information of marker.

The gains in the performance of 3D graphics hardware and rendering systems have not been matched by a corresponding user interface to interact with stereoscopic images. Hence, there is a need of 3D interactive interfaces that can provide continuous, high-freedom, and user friendly interaction for touching the virtual stereo-images of mobile 3D displays or other near-field 3D interactive applications.

II. 3D Interactive Technologies

To fulfill the demand in 3D interaction, many research projects have been devoted to the 3D detection technology.

Manuscript received June 05, 2013; accepted June 24, 2013. Date of publication July 09, 2013; date of current version November 12, 2013.

Y.-P. Huang is with Department of Photonics/Display Institute, National Chiao Tung University, Hsinchu, Taiwan (e-mail: boundshuang@mail.nctu.edu.tw).

G.-Z. Wang and T.-S. Chang are with the Department of Electronics Engineering & Institute of Electronics, National Chiao Tung University, Hsinchu, 300, Taiwan (e-mail: gzwang.ee97g@nctu.edu.tw; tschang@g2.nctu.edu.tw).

T.-H. Chen is with the Cornell University, School of Electrical and Computer Engineering, Ithaca, NY 14853 USA (e-mail: TC345@cornell.edu).

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

Digital Object Identifier 10.1109/JDT.2013.2271633

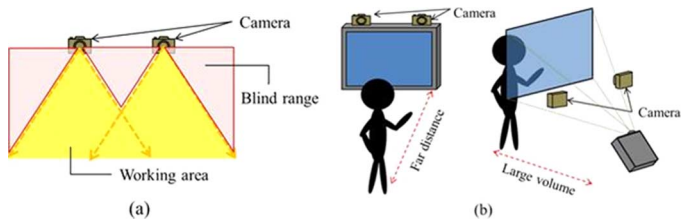


Fig. 1. Limited field of view in cameras causes (a) blind range in continuous interaction space which results in (b) far working distance or large system volume.

Besides 2D position (x, y) on the surface, extra depth information (z) is one of the key parameters needed to be determined. The dominant 3D interactive systems [12] can be classified into three categories depending on their working principle: machine-based, camera-based, and in-cell optical-based. In the following paragraphs, we'll briefly discuss these systems.

A. Machine-Based System

For machine-based systems [13], the user has to wear additional devices to detect motion of user. For instance, Haptic WorkstationTM can detect 6-axis of hand movement through data gloves [14] and render force-feedback on the wrists. Although the machine-based 3D interactive system has the merit of giving force-feedback, it is thought to be inconvenient due to the wearing of additional heavy and bulky devices.

B. Camera-Based System

For camera-based systems [15], 3D information (x, y, z) can be recognized by using various designs of camera technologies. For instance, the popular Wii and Kinect game console [16]–[18] are able to detect relative 3D position by using infrared (IR) cameras corresponding with IR light sources [19].

Another example is WorldViz PPT (Precision Position Tracking) [20]. The optical cameras are mounted in the corners of a room to track the active LED markers. As the object with markers makes its way through the room, the cameras acquire data. Through complex image processing, the cameras' data are converted to 3D positions of the markers.

However, camera-based 3D interactive systems are limited by their field of view which prevent them from detecting objects close to the display and make it difficult for them to perform a continuous interaction space, as shown in Fig. 1. Additionally, a camera-based system requires high resolution data to calculate 3D positions where the resolution is proportional to the size of CCD of the camera, so the form factor impedes the system being integrated into portable devices. Instead of using IR cameras, in-cell optical sensor arrays, which can be integrated within the display pixels to keep the system light and thin was proposed.

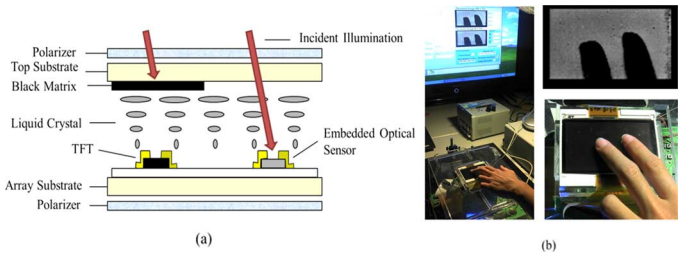


Fig. 2. (a) Schematic structure of embedded optical sensor on TFT substrate and (b) sensed image.

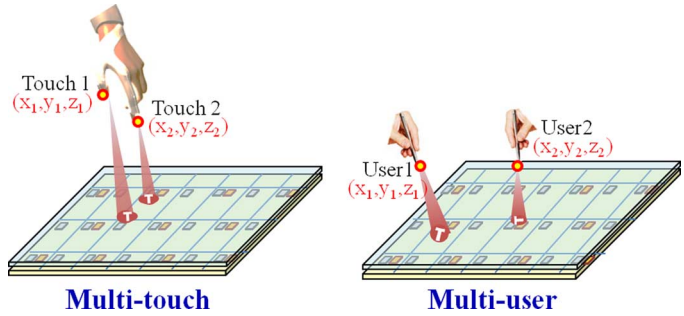


Fig. 3. Concept of Multi-T-mark System for multi-touch operation and multi-user identification.

C. In-Cell Embedded Optical-Based System

Embedded optical sensors onto a TFT array substrate was first proposed by W. D. Boer *et al.* in 2003 [21]. The optical sensors can be fabricated in each sub-pixel using the same process as that of a TFT array. As shown in Fig. 2, where the sensor is not blocked by a black matrix, photo current will be generated when the sensor receives light. Therefore, while putting a hand on the screen, the optical sensors can detect the shadow of a hand, especially the dark fingertips. Consequently, the embedded photo sensing becomes a kind of 2D touch technology.

The 2D optical touch systems are further extended to near-distance 3D interaction. Without touching the surface, the object can be detected by capturing the light incident on a sensor array. One of the systems proposed by Yi *et al.* [22] utilizes backlight reflection to extract the position of fingers on the display surface and of hovering palms. Therefore, multiple touches on a 2D surface and approximate hovering positions can be obtained simultaneously. However, precise depth information (z) cannot be found since the reflection is so weak, thus the captured image will be so blurred that it can only be processed for approximate hover detection.

Another system was proposed by Brown *et al.* [23], where a shielding layer above a PIN photo diode sensor is used to sense the direction of incident light. According to the captured images, the disparity of the light can be calculated and can further determine the 3D position (x, y, z) of the object. However, the result exhibits a limited linear response from 0 to 2 cm, which is still not adequate for 3D virtual touch applications.

By using a high resolution camera approximating high density optical sensors to capture images behind LCDs, Hirsch *et al.* [24] proposed a 3D multi-touch algorithm. By inserting a coded optical mask in front of the panel to capture the multiple images, the algorithm can calculate the depth of an object. However, it requires a very high resolution of sensors yet higher sensor resolution will decrease the aperture ratio of pixels which also increases the loading of readout circuit.

In 2013, the floating touch function of Samsung Galaxy S4 attracts the attention of many consumers and experts. However, this floating touch technology can only provide up to 1.5 cm sensing depth. Additionally, this capacitive floating touch technology cannot detect the accurate 3-axis(x, y, and z) information of object but can support air wave function allows you to swipe between photos by simply waving your hand in front of the screen.

For the prior approaches, the accuracy and working distance are the two general issues. Additionally, some of mobile devices such as ipad will have an urgent need of multi-user operation when the mobile display size becomes more and more large. Furthermore, the multi-touch system can also support zoom in, zoom out, and drag functions for providing friendlier and more intuitive user interfaces.

Therefore, we proposed a Multi-T-mark 3D interactive system which is suitable for the near-field interaction of mobile applications to detect 3-axis(x, y, and z) information of tip of the light-pen and to raise the range of sensing depth. Moreover, not only multi-touch for glove types, but also multi-user operation for light pen types can be achieved in our system, as shown in Fig. 3.

III. Concept of Multi-T-Mark for 3D Interactive

Currently, there are two categories of stereoscopic display systems: glasses type and glasses free type. The most common glasses type is not suitable for mobile applications. Therefore, the glasses free type will be the trend of mobile applications. Then the glasses free type can be divided into lenticular lens-typed and barrier-typed. Lenticular lens-typed is not workable in our proposed light-pen system due to the refractive effect. Therefore, we choose the barrier type stereoscopic display to be our platform for reducing additional optics affecting the functionality of the proposed light-pen system.

For barrier typed embedded optical sensors panels, the bare-finger interaction currently still has some limitations such as insufficient sensor resolution, serious ambient light effect, and inadequate depth sensing range. Therefore, we proposed using IR light-pen to project coded T-marks on the panel. Following, the Multi-T-mark algorithm can be used to detect 3-axis(x, y, and z) information of tip of the light-pen by the orientation information of the tilt angle (θ) and the rotation angle (ϕ). The 2D (x, y) position of tip of the light-pen can be calculated by adding the values of correction factor (Δx and Δy), as shown in Fig. 4(a). The values of correction factor (Δx and Δy) can be deduced from the simple trigonometric function with rotation and tilt angles, as shown in Fig. 4(b). Furthermore, the angular information can also provide more freedom for gaming applications.

Additionally, through properly designed T-marks, multi-user identification and multi-touch applications can be achieved. Although multi-view 3D display also provide different 3D space for different viewer's position. However, this research is focus on detecting the 3-axis(x, y, and z) information of different light-pens simultaneously. Therefore, our objective is to improve the sensing depth and detect the accurate 3-axis(x, y, and

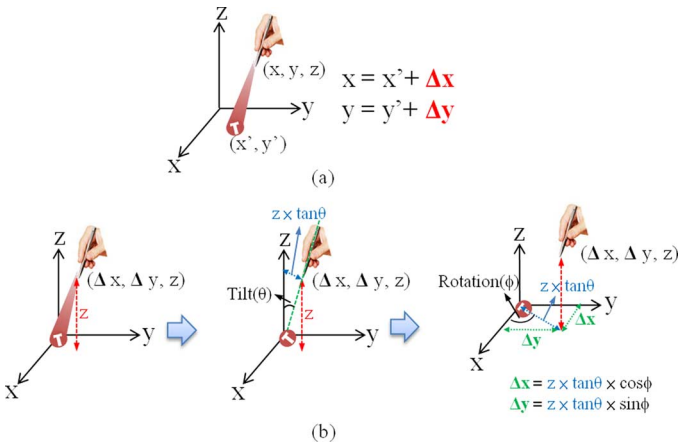


Fig. 4. Concept of Multi-T-mark System for 3-axis(x, y, and z) information of tip of the light-pen (a) definition and (b) determination.

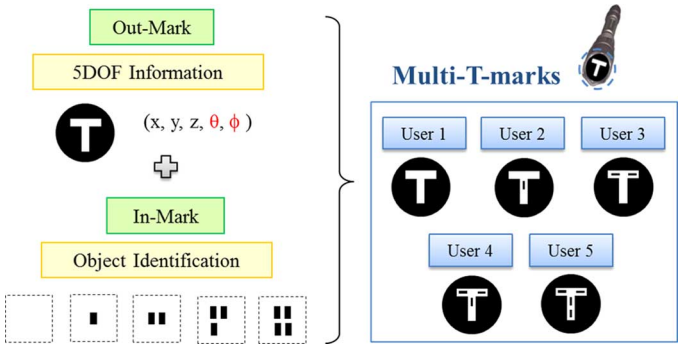


Fig. 5. By combining out-mark and in-marks, multi-T-marks can be used for position detection and user identification.

z) information of multi-touch /multi-user for providing more applications.

A. Mark Design

To recognize or identify the objects, many kinds of Augmented Reality Tags (ARTag) [25] or markers have been developed. Markers with different shapes can be used to separate different users. However, our goal is not only to identify the object but also to find the 5-axis(x, y, z, θ, ϕ) information of the object. Markers with different shapes cannot always provide 5-axis information, and will increasing the calculation complexity in real-time operation. Additionally, the resolution of embedded optical sensors is relatively low compared to that of cameras, thus the mark must be simple. Consequently, the mark comprises two parts: one carries 5-axis information, called out-mark, and the other carries user- identifying information, called in-mark.

A simple yet asymmetrical pattern "T" is utilized as the out-mark to provide not only 3D position (x, y, z) but also the angle of tilt (θ) and rotation (ϕ) which are derived from the patterns directional features. In order to recognize different users, each T-mark is coded with different numbers of inner blocking strips, as illustrated in Fig. 5.

B. Multi-T-Mark Algorithm

A multi-T-mark algorithm [26] is proposed in compliance with the mark design. There are three main parts of the algorithm: the position tracking, the user recognition, and the orientation tracking, to acquire 5-axis information and user identification of the objects. The flow chart of the Multi-T-mark algorithm is illustrated in Fig. 6.

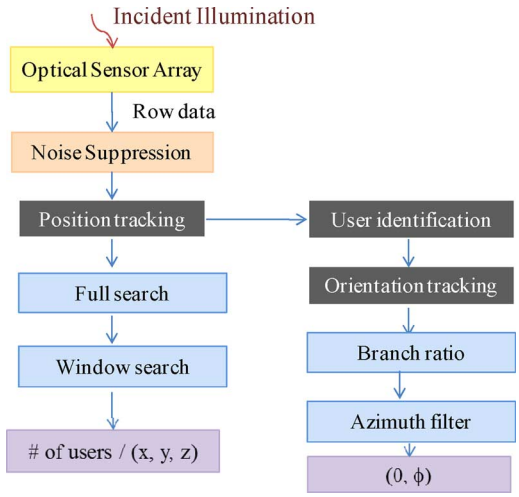


Fig. 6. Flow chart of Multi-T-mark algorithm.

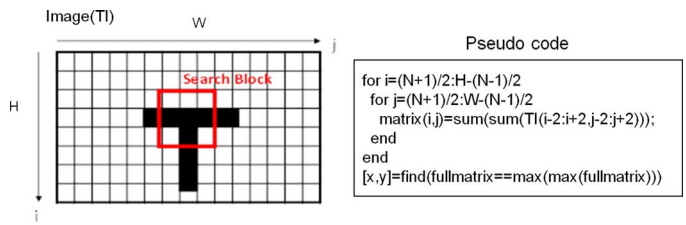


Fig. 7. Full search method for 2D (x, y) position.

Before processing the Multi-T-mark algorithm, de-noise step is necessary for reducing ambient light effect and the impact of the panel defects, such as the system noise and bright spots. After that, the position (x, y, z) of objects will be firstly tracked by using full and window search, and followed by identifying different users. Finally, the orientation (θ, ϕ) of each users will be calculated.

Following are the detail descriptions of the three main parts, position tracking, user identification and orientation tracking, for determining the 5-axis information of the objects and the multi-users.

The first target is to find the 3D position (x, y, z) of the objects. In order to erase the in-marks effect, the in-marks are temporarily filled up by applying image dilation and image erosion processes, and then the labeling process is executed to determine the number of objects.

After knowing the number of objects, a full search is applied to find the 2D (x', y') position of each object, as shown in Fig. 7. The (x', y') position of a T-mark onto the panel can be found by searching the maximum accumulation of N × N search blocks. Then the 2-axis(x and y) information of tip of light-pen can be calculated by adding the values of correction factor (Δx and Δy) to the (x', y') position of a T-mark onto the panel, as shown in Fig. 4.

After knowing the (x, y) location of T-mark onto the panel, the adaptive window which utilizes the variation of T-mark size was applied to determine the depth (z) position of object, as

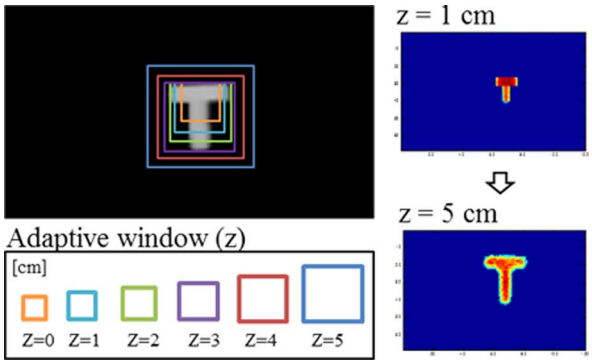


Fig. 8. Adaptive window for depth (z) determination.

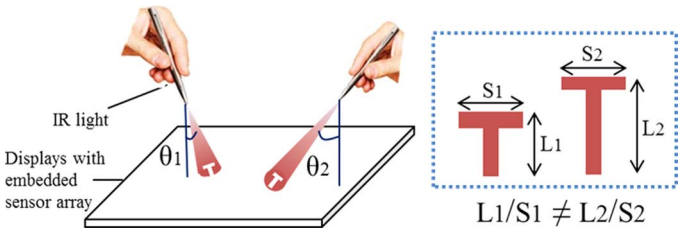


Fig. 9. Branch-ratio method for tilt angle (θ) detection.

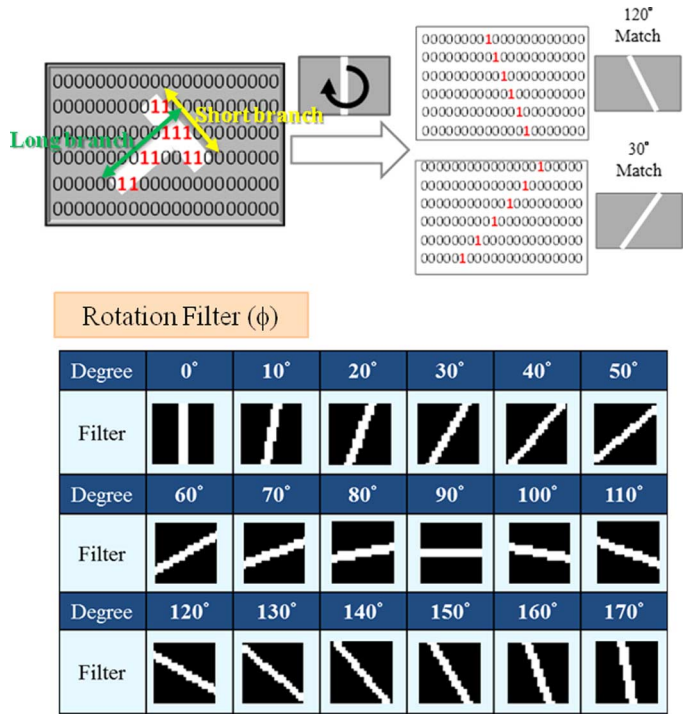


Fig. 10. Rotation filter for rotation angle (ϕ) detection.

shown in Fig. 8. Because the projected IR light has about 25° divergence angle, the size of T-mark will increase as the light-pen moves farther away from the panel. Based on this concept and by using a Look-Up-Table which includes the window size of different heights, the depth (z) position of the object can be obtained.

After knowing the (x, y, z) position, the second part of the Multi-T-mark algorithm is user identification. By normalizing the size of captured T-mark images, an edge filter is applied to extract the in-mark features. For the low resolution of embedded photo sensor, we have to choose the suitable filter for processing in-mark edge. Comparisons of various edge detectors have been made. The Prewitt approximation performs a better edge detection result where the out-marker can be kept, because Prewitt operators can perform better line factor of T marks. Meanwhile, if the in-mark is blurred by light divergence, the 1st derivative process in Prewitt edge filter maintains a better character of the in-mark than other algorithms that perform 2nd derivative in the Laplacian of Gaussian filter or dual threshold in the Roberts and Canny filters. Additionally, from our experiment results, it is no problem in the current experiment platform, even with 68×120 sensor resolutions. Finally, the user identification is achieved by the sorting of the edge accumulation of each mark window.

In the last part, the orientation information, the angles of tilt (θ) and rotation (ϕ), can be acquired by the Branch-ratio method and the Azimuth filter, respectively. The Branch-ratio method is utilized to determine the tilt angle (θ) by the ratio of the long-branch divided by the short-branch of the out- mark, as illustrated in Fig. 9. As the objects are tilted, the length of the bars changes according to the tilt angle. Therefore, the tilt angle (θ) can be calculated.

Finally, an Azimuth filter which includes ration filter and branch filter is applied to acquire the rotation angle (ϕ). Rotation filter is utilized to acquire the orientation angle. By considering the particle implementation in real-time, a look-up table for the rotation filter with approximate ratio was established instead of calculating the absolute rotation angle, as shown in Fig. 10. Once the bar of the T-mark matches the filter angle, it is recorded as the accordant rotation filter which reveals the highest similarity of the T-mark to the filter. However, without knowing the orientation of the short branch, the precise rotation angle (ϕ) cannot be identified. For the rotation filter, it actually indicate two angles: θ and $\theta + 180°$ due to the same orientation of their long branches. Hence, it is essential to define the orientation of a short branch by utilizing the branch filter. Branch filter is used for identifying the orientation of a short branch and acquired the rotation angle. The watershed of a branch filter is perpendicular to the according rotation filter, as shown in Fig. 11. By knowing the 2D coordinate (x and y) and the rotation angle of T-mark, the accordant branch filters, up and down, are applied for branch identification. If an accumulation result reveals a larger up filter, the short branch locates in the up filter. Hence the rotation angle (θ) equals to the angle of accordant rotation filter. On the other hand, if the accumulation of a down filter is larger than up filter, the short branch locates in the down filter. Hence there should be a 180° modification, where the rotation angle is $\theta + 180°$.

Accordingly, by using the Multi-T-mark algorithm, the 3-axis information (x, y, and z) of tip of the different light-pens can be obtained to further extend the freedom in 3D interaction applications.

IV. Experiment Results

A prototype 3D interactive system which can concurrently sense 3D virtual touches with 3-axis information (x, y, and z) and multi-user identification was built, as shown in Fig. 12, with a 4" LCD equipped with IR sensors integrated onto the TFT substrate.

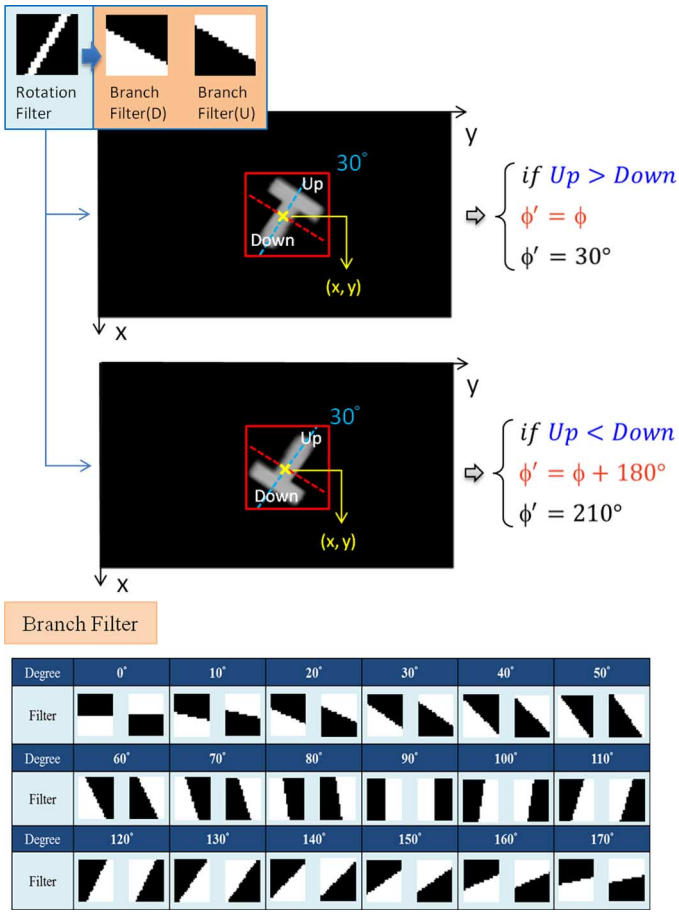


Fig. 11. Branch filter for rotation angle (ϕ) definition.

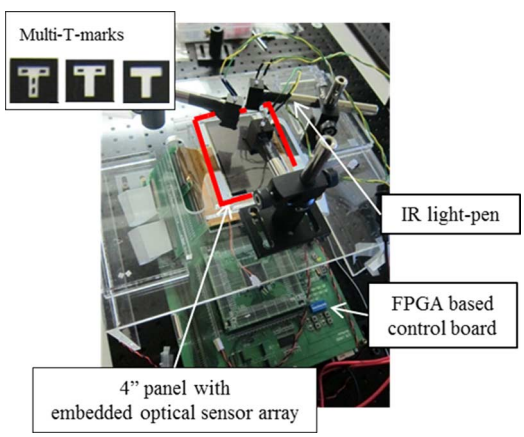


Fig. 12. Prototype of the 3D multi-interactive system with 4" sensible LCD notarizes the Multi-T-mark interaction. The video link: http://www.youtube.com/watch?v=NgomHrS9r4Y&feature=youtu.be The video (3D Virtual Touch) link: http://youtu.be/ipJeMAdZVCI.

The sensor resolution is 68×120, which is one fourth of the image pixel resolution. The FPGA-based control unit drives the IR sensor array to operate at 30 Hz to detect the incident IR light. An 850 nm IR LED with 25° divergence angle was used as the input light source with an operating voltage of 3 V and operating current of 0.15 A. The IR light pens can project T-marks for demonstrating the Multi-T-mark algorithm.

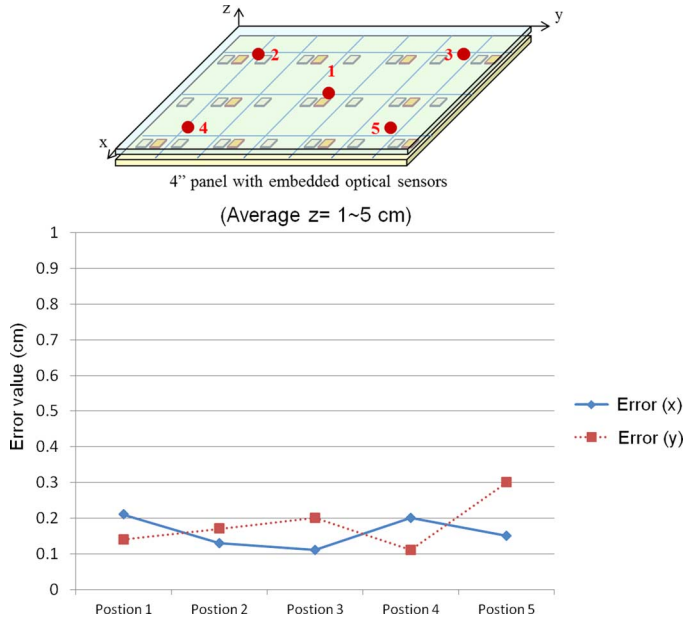
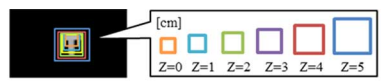


Fig. 13. Result of 2-axis (x and y) measurements.



z [cm]	0	1	2	3	4	5
Window size [pixel]	11	14	19	23	27	31

Fig. 14. Adaptive window sizes for depth value detection.

A. Resolution in x' and y'

In 2-axis (x' and y') detection, we set the light-pen with zero rotation angle and zero tilt angle to raise from 1 cm to 5 cm in different positions (position 1 to position 5), as shown in Fig. 13. The average error of 2-axis (x' and y') value was smaller than 3 mm. Furthermore, the maximum error in the (x', y')-plane were 3 mm. Therefore, the experiment result indicates the detected coordinates of X′/Y′ matched the real value. Further to discussions at the 2-axis (x and y) detection with nonzero rotation angle and nonzero tilt angle, we can see that the 2D (x, y) position of tip of the light-pen can be calculated by adding the values of correction factor (Δx and Δy) to the (x', y') position of a T-mark onto the panel. Additionally, the values of correction factor (Δx and Δy) are depends on the rotation and tilt angles. Therefore, the rotation and tilt angles will affect the 2-axis (x and y) detection due to the values of correction factor (Δx and Δy).

B. Resolution in Depth (z)

In our system, the resolution in the z-axis is determined by the adaptive window we designed. Then the optimized adaptive window size is shown in Fig. 14.

To further check the depth detection in the z axis, three objects with various T-marks were placed at different heights from 0 to 5 cm. To demonstrate the continuous transfer in real conditions, a 0.5 cm interval was set in the experiments. As show in Fig. 15, the average depth detection of the three T-marks is almost linear, and is close to the ground truth (Actual depth) value.

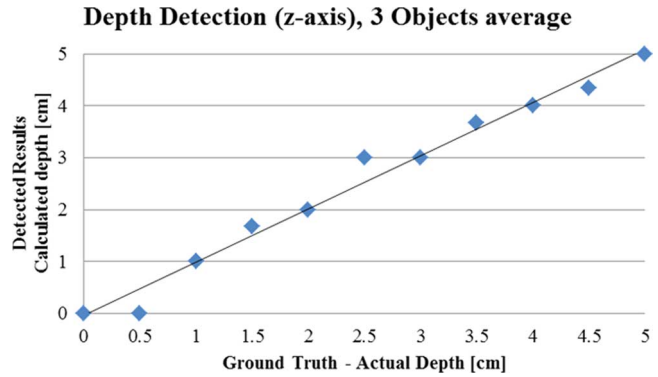


Fig. 15. Result of depth measurements.

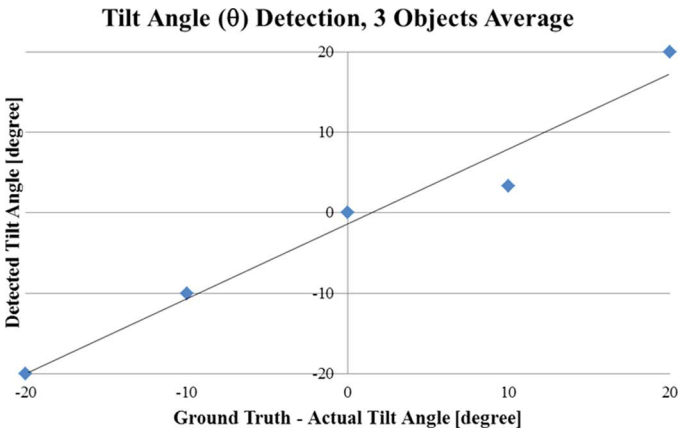


Fig. 16. Result of tilt angle measurements.

In our study, the proposed systems are suitable for near-field interaction of mobile applications. Therefore, the limited working space (0–5 cm) in depth should be enough. Furthermore, the resolution in depth could be further improved by increasing the embedded sensor resolution.

C. Resolution in Orientation (θ, ϕ)

In tilt angle (θ) detection, we set the resolution to 10 degrees, and the detection range from -20° to $+20^\circ$. The tilt angle range is limited by the panel size of our prototype and the serious mark distortion over 30°. By subtracting each branch of the T-mark, the rotation angle can be calculated by the branch ratio. The result is shown in Fig. 16.

Meanwhile, the resolution of the rotation angle (ϕ) was verified. In the experiments, we calibrated the rotation angle to zero degrees as the vertical branch of T is parallel to x-axis. A 360° rotation was made with 5° intervals to verify the fidelity of continuous rotation about z-axis. The result, as shown in Fig. 17, reveals a high match with the ground truth.

D. Multi-User Identification Experiments

To prove the feasibility of multi-user identification, IR light pens with multi-T-marks were placed at different positions with different orientations, as shown in Fig. 18(a). The maximum of three light pens, which represents three users, can interact with our prototype simultaneously. The experiment results shown in Fig. 18(b) and (c) indicate the ground truth and the detected

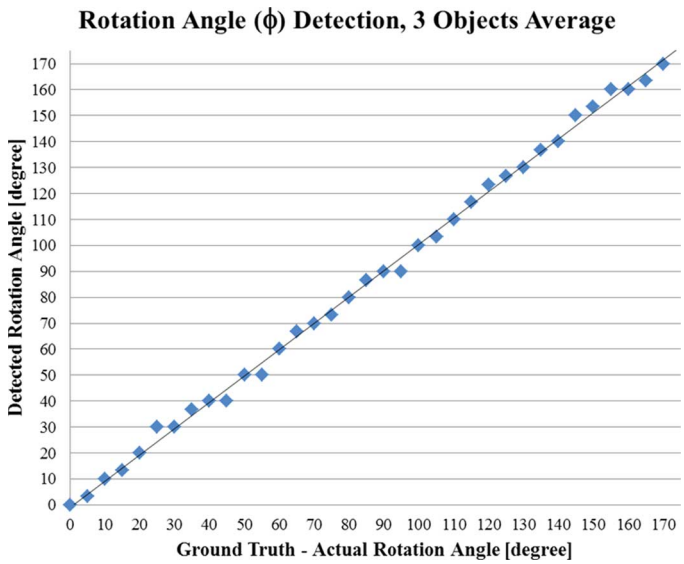
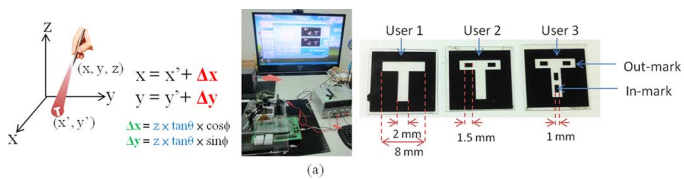


Fig. 17. Result of rotation angle measurements.



(a)

Ground Truth		x (cm)	y (cm)	z (cm)	Tilt(θ°)	Rotation(ϕ°)
Test 1	User 1	0.92	1.28	4	0	0
	User 2	1.79	3.51	0	0	10
	User 3	1.24	2.27	2	0	170
Test 2	User 1	1.06	0.99	2	0	90
	User 2	1.39	3.11	2	10	60
	User 3	1.17	2.05	2	20	10

(b)

Detected Results		x (cm)	y (cm)	z (cm)	Tilt(θ°)	Rotation(ϕ°)	x' (cm)	y' (cm)	Δx (cm)	Δy (cm)
Test 1	User 1	1.06	1.35	4	0	0	1.06	1.35	0	0
	User 2	1.9	3.55	0	0	10	1.9	3.55	0	0
	User 3	1.35	2.34	2	0	170	1.35	2.34	0	0
Test 2	User 1	1.1	0.99	2	0	90	1.1	0.99	0	0
	User 2	1.39	3.26	2	10	60	1.21	2.96	0.18	0.3
	User 3	1.24	2.12	2	20	10	0.52	1.99	0.72	0.13

(c)

Fig. 18. Experiment results of 5-axis detection and 3-user recognition.

results. From the experiment results, we can see that the proposed 5-axis detection has high accuracy. Only some of the 2D (x, y) coordinates detections have a few centimeter displacement, which are about 2 mm only thus can be ignored. The calculated results also evinces that robust user identification can be achieved by the in-mark patterns.

V. Conclusion

To fulfill the demands of "touching" stereo-images on mobile 3D display with multi-user/multi-touch function, the embedded optical sensor based structure is chosen by virtue of its thin-form factor, camera free, and apposite working range that is suitable for mobile display integration and creates a continuous

working space from the surface to near distance. We have shown that an embedded optical sensor panel in collaboration with the Multi-T-mark system can be used as a 3D input device for accurate detection of the 3-axis information (x, y, and z) of the tip of the light-pen and achieve accurate user identification. The patterned infrared light sources and the Multi-T-mark algorithm are proposed where the out-mark is utilized to determine the 5-axis information (x, y, z, θ, ϕ) of the tip of the light-pen, and the in-mark further renders multi-user identification. The experimental results exhibit precise 2D position (x, y) detection and a linear response in depth value (z) from 0 to 5 cm. Now most of mobile 3D display can only provide shallow depth. That's because glasses-free 3D technology still has cross-talk which limited the displayed image depth. Therefore, the proposed system which can provide 0~ 5 cm working range is suitable for the near-field interaction of mobile applications. The orientation information (θ, ϕ) is achieved in 10 degree steps. For further enhancing the resolution of 5-axis, we can increase the resolution of photo sensors in the future. Finally, the 3D interactive technology was successfully implemented on a 4" embedded optical sensor display.

REFERENCES

[1] K. Bredies, N. A. Mann, J. Ahrens, M. Geier, S. Spors, and M. Nischt, "The multi-touch soundscape renderer," in *AVI '08: Proc. Working Conf. on Adv. Visual Interfaces*, 2008, pp. 466–469.

[2] W. Mphepö, Y.-P. Huang, and H.-P. D. Shieh, "Enhancing the brightness of parallax barrier based 3D flat panel mobile displays without compromising power consumption," *J. Display Technol.*, vol. 6, no. 2, pp. 60–62, Feb. 2010.

[3] M. Tsuboi, S. Kimura, Y. Takaki, and T. Horikoshi, "Design conditions for attractive reality in mobile-type 3-D display," *J. Soc. Inf. Disp.*, vol. 18, pp. 698–703, 2010.

[4] Y. P. Huang, L. Y. Liao, and C. W. Chen, "2-D/3-D switchable autostereoscopic display with multi-electrically driven liquid-crystal (MeD-LC) lenses," *J. Soc. Inf. Disp.*, vol. 18, pp. 642–646, 2010.

[5] K. M. Chen, S. Gauza, H. Xianyu, and S.-T. Wu, "Submillisecond gray-level response time of polymer-stablized blue-phase liquid crystal," *J. Display Technol.*, vol. 6, no. 2, pp. 49–51, Feb. 2010.

[6] J. Nakamura, T. Takahashi, C. W. Chen, Y. P. Huang, and Y. Takaki, "Increase in longitudinal viewing freedom of reduced-view super multi-view display," *J. Soc. Inf. Disp.*, vol. 4, pp. 228–234, 2012.

[7] Y. P. Huang, C. W. Chen, T. C. Shen, and J. F. Huang, "Autostereoscopic 3D display with scanning multi-electrode driven liquid crystal (MeD-LC) lens," *J. 3D Res.*, vol. 1, pp. 39–42, 2010.

[8] S. Jumisko-Pyykk, M. Weitzel, and D. Strohmeier, "Designing for user experience: What to expect from mobile 3D TV and video?," in *Proc. 1st Int. Conf. on Designing Interactive User Experiences for TV and Video*, 2008, pp. 183–192.

[9] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. E77-D, no. 12, pp. 1–15, 1994.

[10] D. W. F. van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *IJVR*, vol. 9, no. 2, pp. 1–20, 2010.

[11] H. Kato and M. Billinghurst, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," *IWAR*, pp. 85–94, 1999.

[12] C. Hand, "A survey of 3D interaction techniques," *Comput. Graph. Forum*, vol. 16, pp. 269–281, 1997.

[13] R. Ott, M. Gutiérrez, D. Thalmann, and F. Vexo, "Advanced virtual reality technologies for surveillance and security applications," in *Proc. 2006 ACM Int. Conf. on Virtual Reality Continuum and its Appl.*, 2006, pp. 163–170.

[14] K. N. Tarchanidis and J. N. Lygouras, "Data glove with a force sensor," *IEEE Trans. Instrum. Meas.*, vol. 52, no. 3, pp. 984–989, Mar. 2003.

[15] A. Olwal, S. DiVerdi, N. Candussi, I. Rakkolainen, and T. Hollerer, "An immaterial, dual-sided display system with 3D interaction," in *Virtual Reality Conf.*, 2006, pp. 279–280.

[16] T. Schlömer, B. Poppinga, N. Henze, and S. Bol, "Gesture recognition with a Wii controller," in *Proc. 2nd Int. Conf. on Tangible and Embedded Interaction*, 2008, pp. 11–14.

[17] 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. on Comput.-Human Interaction: Entertaining User Interfaces*, 2007, pp. 231–234.

[18] B. 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.

[19] T. Shiratori and J. K. Hodgins, "Accelerometer-based user interfaces for the control of a physically simulated character," *ACM SIGGRAPH Asia*, p. 123, 2008.

[20] 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.

[21] 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.

[22] K. Yi, C. Choi, S. Suh, B. Yoo, J. J. Han, D. Park, and C. Kim, "45.2: Distinguished paper: Novel LCD display with a sensible backlight," in *SID Symp. Dig. Tech. Papers*, 2010, vol. 41, pp. 673–676.

[23] C. 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.

[24] M. Hirsch, D. Lanman, H. Holtzman, and R. Raskar, "BiDi screen: A thin, depth-sensing LCD for 3D interaction using light fields," *ACM SIGGRAPH Asia*, pp. 62–65, 2009.

[25] M. Fiala, "ARTag, A fiducial marker system using digital techniques," *Comput. Vision and Pattern Recogn.*, vol. 2, pp. 590–596, 2005.

[26] Y. P. Huang, G. Z. Wang, S. Y. Huang, Y. H. Tai, and T. H. Chen, "Camera free 3-dimensional virtual touch display with multi-user identification," in *19th IEEE Int. Conf. on Image Process. (ICIP)*, 2012, pp. 1965–1968.

Yi-Pai Huang received the B.S. degree from National Cheng Kung University in 1999, and the Ph.D. degree in electro-optical engineering at the National Chiao Tung University in 2004.

He is currently an full-time associate professor in the department of photonics & display institute, National Chiao Tung University, Taiwan. Also, he was a visiting associate professor at Cornell University, Ithaca, NY, USA, from 2011 to 2012. His expertise includes 3D Display and Interactive Technologies, Display Optics and Color science, Micro-optics. In the above-mentioned research, he have so far published more than 100 international journal and conference papers (including the 50 SID Conference Papers and 7 invited talks), and have obtained 25 granted patents, with another 48 patents currently publicly available.

Dr. Huang is the secretary general of SID Taipei Chapter, and Chair of APV program sub-committee, SID. In addition, He had three times received the SID's distinguished paper award (2001, 2004, 2009). Other important awards include 2011 Taiwan National Award of Academia Inventor, 2010 Advantech Young Professor Award, 2009 Journal-SID: Best paper of the Year Award, and 2005 Golden Dissertation Award of Acer Foundation.

Guo-Zhen Wang received the M.S. degrees from Display Institute, National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 2008, and is currently working toward the Ph.D. degree at the Department of Electronics Engineering, National Chiao Tung University (NCTU).

His current research is to develop 3D interaction systems, especially in image processing, and computer architecture technologies.

Tian-Sheuanm Chang (S'93–M'06–SM'07) received the 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 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 in IMEC, Belgium. His current research interests include system-on-a-chip design, VLSI signal processing, and computer architecture.

Dr. Chang has received the Excellent Young Electrical Engineer from 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 an organizing committee or technical program committee member. He is current an Editorial Board Member of IEEE TRANSACTIONS OF CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.

Tsuhan Chen received the B.S. degree from National Taiwan University and the M.S. and Ph.D. degrees from Caltech, Pasadena, CA, USA, all in electrical engineering.

After working for Bell Labs for several years, he joined the ECE faculty at Carnegie Mellon University, Pittsburgh, PA, USA, in 1997. There, in addition to research and teaching responsibilities, he has served as associate department head of ECE and co-director of the Industrial Technology Research Institute (ITRI) Laboratory at Carnegie Mellon, a collaborative research program with ITRI in Taiwan. In January 2009, he was chosen to become the director of the School of Electrical and Computer Engineering at Cornell University, Ithaca, NY, USA, after an extensive national search.