Compact and High Resolution Virtual Mouse Using Lens Array and Light Sensor

Zong Qin[a], Yu-Cheng Chang[a], Yu-Jie Su[a], Yi-Pai Huang*[a], Han-Ping David Shieh[a]
[a]Department of Photonics & Institute of Electro-Optical Engineering & Display Institute, National Chiao Tung University, 300, Hsinchu, Taiwan, ROC +886-3-5712121 ext. 59210;

ABSTRACT

Virtual mouse based on IR source, lens array and light sensor was designed and implemented. Optical architecture including lens amount, lens pitch, baseline length, sensor length, lens-sensor gap, focal length *etc.* was carefully designed to achieve low detective error, high resolution, and simultaneously, compact system volume. System volume is 3.1mm (thickness) × 4.5mm (length) × 2, which is much smaller than that of camera-based device. Relative detective error of 0.41mm and minimum resolution of 26ppi were verified in experiments, so that it can replace conventional touchpad/touchscreen. If system thickness is eased to 20mm, resolution higher than 200ppi can be achieved to replace real mouse.

Keywords: Virtual mouse, lens array, position detection, light sensor, input device

1. INTRODUCTION

The concept of virtual mouse refers to a way users move their finger in free space instead of real mouse or on-device touch for notebook, tablet or other consumer electronics, as shown in Figure 1[1-7]. Virtual mouse device provides positional information of human finger to consumer electronics to play the role that is originally played by real mouse or touchpad/touchscreen. Therefore, with the aid of virtual mouse, users can get rid of real mouse when they go out with their notebook. On the other hand, moving finger in free space beside the device is more comfortable than being restricted in the region of touchpad or touchscreen. Even, 3-dimensional (3D) positional information can be detected via virtual mouse to provide more possibilities for applications, like gesture control. Therefore, it is quite attractive and promising to develop practical virtual mouse to provide convenience to users and augment interactive input device.

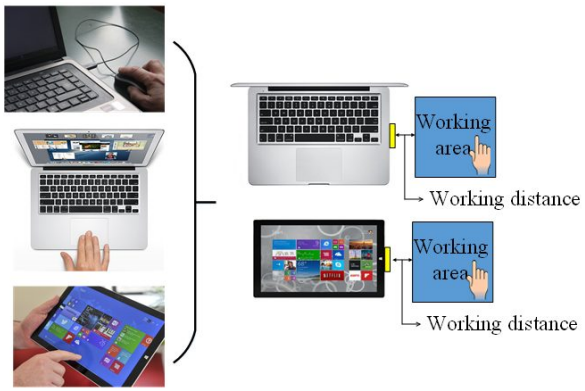

Figure 1. Concept of virtual mouse, and definition of working area and working distance

Key specifications that can make virtual mouse practicably replace real mouse or touchpad/touchscreen mainly include: (1) Working area should be beside the device for users to comfortably place their finger in, and with appropriate size and working distance, as shown in Figure 1. In this paper, all the design will be implemented for working area size of 10cm by 10cm and working distance of 12cm. (2) Resolution of virtual mouse should be no lower than that of current device. Specifically, to replace touchpad/touchscreen or opto-mechanical mouse, the minimum requirement of resolution in the working area is 25ppi or 200pi respectively. (3) Accuracy should be sufficient to avoid perceivable detective error. By considering that it is relative but not absolute positional information that is utilized by mouse device, distinguishable point number that standard deviation of detective error corresponds to should be no larger than 0.5. (4) Volume should

Three-Dimensional Imaging, Visualization, and Display 2016, edited by Bahram Javidi, Jung-Young Son, Proc. of SPIE
Vol. 9867, 986718 · © 2016 SPIE · CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2229446

be as small as possible since virtual mouse will be assembled in consumer electronics. Volume of millimeter scale is recommended. (5) System architecture and detection algorithm should be simple enough for fast response and low cost. Moreover, detection should be stable against the interference of background.

Technologies that can implement position detection, like virtual mouse, can be classified into two types, as camera-based and emitter-sensor-based[7-8]. Generally, camera-based technology is better in high resolution and anti-interference[9-11] while source-sensor-based technology is better in response time and compact volume[12-15]. However, as mentioned above, high resolution, anti-interference, fast response time, compact volume, along with acceptable cost are all considerably concerned for practical virtual mouse. It is quite difficult to find a current technology that can achieve these features simultaneously.

In this paper, a virtual mouse prototype based on infrared radiation (IR) source, lens array, and linear light sensor was designed and implemented. Compact system volume of 3.1mm (thickness) × 4.5mm (length) × 2 was achieved, which is much smaller than that of camera-based device. Two groups of lens arrays designed in house covering on two linear light sensors with pixel size of 2μm could achieve resolution higher than 25ppi in a working area with size of 10cm by 10cm and working distance of 12cm. Experiments verified expected function by using proposed low-computation detection algorithm. Standard deviation of detective error was as low as 0.41mm, which corresponds to only 0.40 distinguishable points that cannot be perceived by users. Measured minimum resolution in the working area was 26ppi corresponding to 2μm pixel size. Sufficient accuracy and resolution, along with compact volume, low-computation detection algorithm, anti-interference feature of IR source and simple system architecture reveal that proposed virtual mouse can practicably replace conventional touchpad or touchscreen. If thickness of the system is relaxed to 20mm, resolution in the working area can be higher than 200ppi, so that it can replace real mouse with acceptable system volume.

2. OPTICAL ARCHITECTURE

To detect finger position, lens arrays covering on light sensors were adopted in our study, as shown in Figure 2. An IR source was placed in the middle of the two lens arrays. A finger moving in working area reflects IR light, then convex lenses converge light and generate multiple focused images on the light sensors in compliance with imaging law of lens. In consequence, as long as we capture the signal on the light sensor and recognize separated images, position of the finger can be known on the chief ray of imaging path. If more than one images are captured, intersection points of multiple imaging paths can determine the finger position.

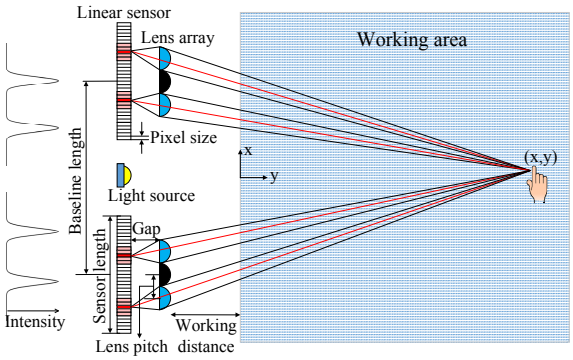

Figure 2. Optical architecture of proposed virtual mouse, including light sensors and lens arrays

This optical architecture can realize finger position detection theoretically. However, to achieve sufficient resolution and accuracy for practical virtual mouse, parameters including lens amount, lens pitch, baseline length, sensor length, sensor-lens gap, focal length, etc., need to be carefully determined. In case that sensor's pixel size is 2μm, following discussion will introduce how to determine these parameters for working area size of 10cm by 10cm and working distance of 12cm.

A. Sufficient resolution

For a mouse, resolution is the key specification that determines its sensibility and comfortability. In our architecture, we can distinguish different finger positions only if captured signals on the sensors are different.

As shown in Figure 2, six lenses in two lens arrays were adopted in our system. Two lenses in the middle of each lens array were painted black, hence four images can be finally captured. (The way we set lens arrays will be explained in

Part B of this section). Pinhole model was used to approximate the system. As shown in Figure 3, the lens that is the farthest from the object can reflect object movement the most sensitively. Therefore, for the farthest lens, image position M with respect to system's axis can be calculated by Equation (1) where BL is baseline length, d is lens pitch, G is sensor-lens gap and (X,Y) is finger position.

Obviously, image position can reflect lateral object movement much more sensitively than vertical movement, hence vertical resolution RV was considered here to investigate system resolution. Variation of Y corresponding to size of a pixel is a distinguishable vertical object movement. Here negative variation of Y, as object moving towards the lens, was considered. Then vertical resolution can be calculated by Equation (2) where μ is pixel size. From Equation (2), the smallest vertical resolution occurs at X=0 and Y=22cm, hence vertical resolution here determines system resolution.

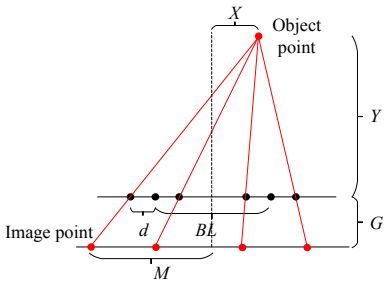

Figure 3. Pinhole model for resolution calculation

$$M = (BL/2 + d + X)\cdot(1 + G/\mathrm{Y}) - X \tag{1}$$

$$R_V\,(\mathrm{ppi}) = 25.4 \times [\frac{G\cdot(BL/2 + d + X)}{\mu\cdot Y^2} + \frac{1}{Y}] \tag{2}$$

If we want our virtual mouse to replace conventional touchpad or touchscreen, 25ppi is expected. Moreover, sensor-lens gap G should be small in consideration of small system thickness. In light of this, a possible solution for R_v=25ppi@(X=0, Y=22cm) is BL=60mm, G=3mm and d=1mm. Then, for Y varying from 12cm to 22cm, vertical resolutions corresponding to X=0, X=2.5cm and X=5cm are shown in Figure 4(a). Additionally, Figure 4(b) shows resolutions corresponding to G eased to 25mm. In this case, resolution keeping higher than 200ppi can lead our virtual mouse to replace real mouse.

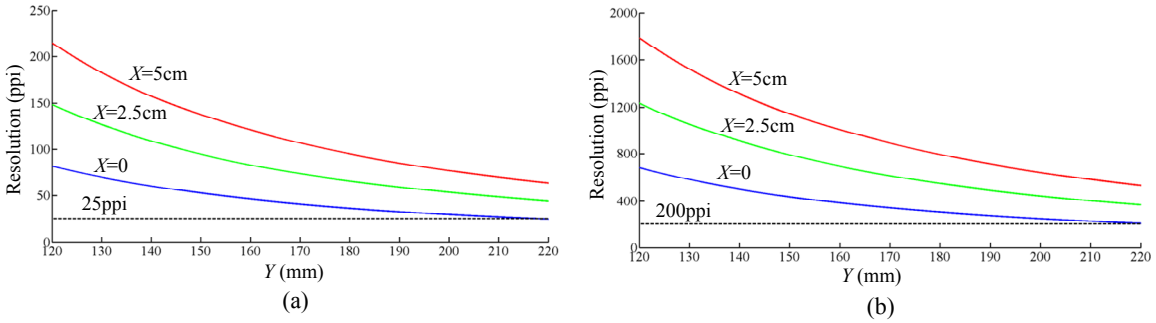

Figure 4. Vertical resolution varying with X from 12cm to 22cm: (a) Sensor-lens gap G=3mm; (b) Sensor-lens gap G=25mm.

B. Separated and clear images

After determining sensor-lens gap, baseline length and lens pitch, we needed to further determine lens amount, focal length and sensor length to guarantee separated and clear images when the finger is located at any position of the working area, *i.e.*, any field of view. Due to considerable aberration of single-element wide-angle lens, produced image is distorted[16]. On the other hand, aspheric lens or multi-element lens system, which can effectively suppress aberration, will not be utilized in our system because they lead to significantly higher cost and larger thickness, which is concerned more than system length is for compact volume. Therefore, more than two images are expected to suppress detective error caused by aberration. Finally, six lenses in two lens arrays with baseline length of 60mm that was determined before were adopted.

By considering that produced image will be blurred due to depth of focus (DOF) while the finger moves in the working area, we should find out a focal length that achieves images as small as possible over the whole working area to prevent image superposition. Lens pitch (lens aperture) had been determined in advance, then the ratio of blurred image's diameter to lens aperture was adopted to quantitatively describe defocusing, as defocusing coefficient D, as shown in Figure 5(a)[16]. To determine the optimum focal length, D varying along y-axis direction corresponding to different focal lengths is shown in Figure 6(b), from which focal length of 2.9mm was selected as the optimum one because it can achieve smaller value of D in the working area than other focal lengths can.

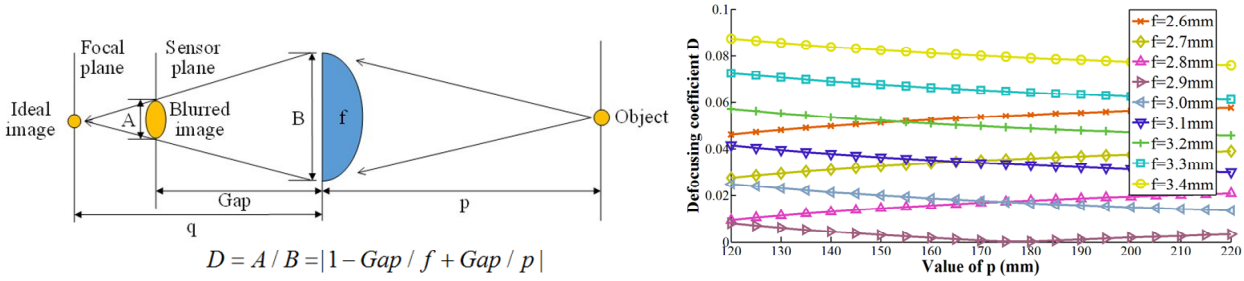

Figure 5. (a) Calculation schematic of defocusing coefficient D (b) Defocusing coefficient D varying along y-axis direction in the working area (p=12cm~22cm) corresponding to focal lengths from 2.6mm to 3.4mm

Next, after focal length was determined, to guarantee separated images, the two lenses in the middle of each lens array were painted black, as shown in Figure 2. For verification, some fundamental optical formulas were used to calculate produced images on the two sensors. Based on Gaussian Optics, object distance p, image distance q, and focal length f obey $1/p+1/q=1/f$[16], hence focal plane can be acquired. Then on the sensor plane, images of a 10cm-long finger can be calculated with the aid of simple geometric principle, as illustrated in Figure 6. Six representative finger positions shown in Figure 7, i.e. six different fields of view, were investigated. By considering that Gaussian Optics may introduce a little error, LightTools simulation were directly used for verification. A 100mm-long cylindrical surface source with 5mm radius was constructed to simulate the finger. Figure 8 shows simulation results. It can be seen that four clearly separated images can be always produced on two sensors with sufficient space. The smallest space is 0.54mm.

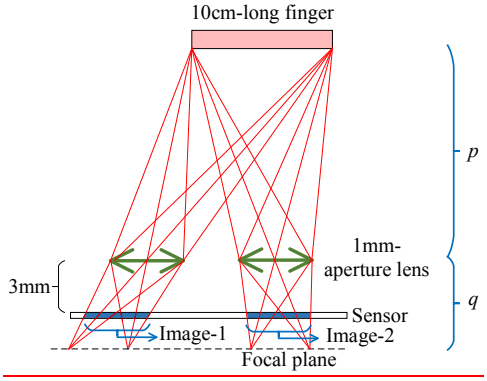

Figure 6. Schematic of calculating finger's image on the two sensors

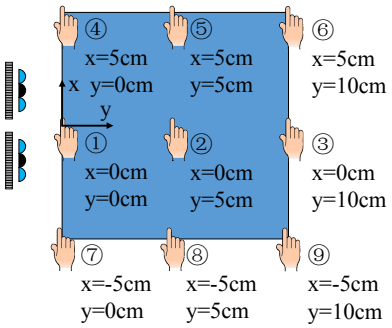

Figure 7. Six representative finger positions for investigation

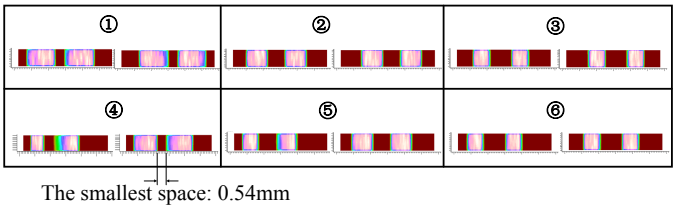

Figure 8. Simulated irradiance distributions on the two light sensors corresponding to six representative finger positions

In conclusion, six lenses with pitch of 1mm and focal length of 2.9mm, including two painted black, were arranged in two groups of lens arrays with baseline length of 60mm. The gap between lens arrays and light sensors is 3mm. If lens arrays are made up of PMMA material, overall thickness comprising the gap and lens arrays is 3.10mm. Each sensor is 4.5mm long with pixel size of 2μm.

3. DETECTION ALGORITHM

After expected signals on light sensors are captured, detection algorithm is needed to solve finger position. Besides sufficient detective accuracy, computations of the algorithm should be as simple as possible in consideration of the requirement of fast response and low cost. Here a detection algorithm with low computation amount was proposed, which comprises two simple procedures:

A. Pre-process

Firstly, noise in captured raw signals will be reduced with a denoise filter, like average filter. Then a threshold will be adopted to eliminate low-luminosity stray light that may be caused by background. Finally, processed signal will be binarized, and clear light pattern(s) with definite edges will be obtained on each sensor. Those complete light patterns whose two edges are both located on a sensor will be considered for subsequent process, while those incomplete ones will be ignored. Moreover, between two light patterns on a sensor (despite complete or incomplete ones), there will always be a dark pattern. It will be also involved in subsequent process. In fact, appropriate design of optical architecture for specific working area can guarantee two complete light patterns on each sensor in most cases. Even though the finger is located in close corners, at least one complete light pattern along with an incomplete one can still be obtained, then detection algorithm can be implemented successfully.

B. Solving of finger position

After pre-process, at least two and at most four light patterns along with two dark patterns will have been obtained on two sensors. Then center of each pattern will be determined to be the position of an image generated by corresponding lens, including dark patterns corresponding to painted lenses. In our set-up, sensors were directly covered on the bottom of lens arrays, as shown in Figure 9, hence thick lens immersed in air on both sides was considered. Its node points coincide with principle points, as N_1 and N_2 in Figure 9. Based on graphical method of thick lens imaging[16], chief ray can be drawn from imaging point to object space. Then several intersection points can be obtained between pair-wise rays. Finally, average position of these intersection points will be determined to be the finger position. The whole detection algorithm was illustrated in Figure 9. However, node points are just first order property of an optical system, and difference exists between it and the actual one. Therefore, as mentioned before, it is necessary to adopt multiple lenses to suppress error. The accuracy of this method will be experimentally verified in the next Section.

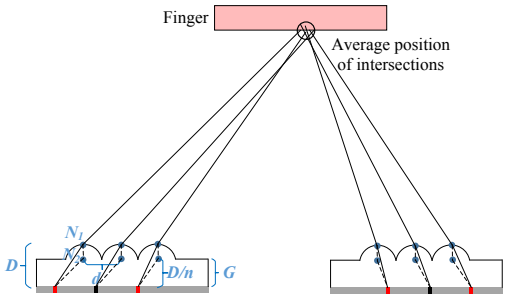

Figure 9. Schema of finger position detection algorithm where red and black marks denote central points of captured patterns, blue marks denote node (principle) points of single lens. D is lens thickness, n is lens material's refraction index, G is sensor-lens gap and d is lens pitch.

4. EXPERIMENTS AND DISCUSSION

Figure 10 shows the experimental architecture, in which lens arrays were actually fabricated by Coretronic Corp. For the moment, we were not able to get appropriate linear light sensor with 2μm, then two planar CMOS sensors with pixel size of 4μm and dimension of 4.5mm×3mm from commercial webcams were used instead. In this way, resolution of this prototype will be half of design value correspondingly. For this verifying prototype, white LEDs were used instead of IR source. In a darkroom, background is far enough from the finger and reflectance of the finger is much larger than the track. Therefore, only finger can be captured, like the situation of IR source. If the experiment succeeds to verify expected function, IR source can be further adopted to make this prototype an actual product with feature of anti-interference-free. A 10cm-long fake finger assembled on a 2-dimensional (2D) track was used to simulate moving human finger. Surface of the fake finger was diffusely reflective. In addition, denoise and binarization will be implemented for captured images. Hence captured brightness can be always uniform and source placement cannot significantly affect imaging system. Those six finger positions in Figure 7 were also experimentally investigated. Performances of relative error and resolution were expected to be verified.

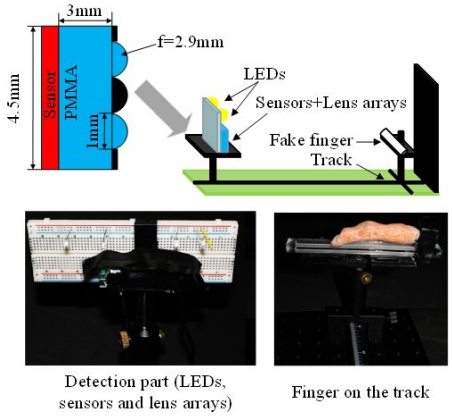

Figure 10. Experimental architecture

After locating center of the fake finger at the six positions, 6×2 raw signals were captured, as shown in Figure 11(a). In pre-process, an average filter with filter size of 5 was used to reduce noise, and then threshold of 10% full-on grayscale was used to cut off stray light, which resulted in clear binary images, as shown in Figure 11(b). By considering that the planar sensors used in the experiments were just replacement of the linear sensors in design, the row of pixels that vertically faced the lens arrays were extracted, and 1-dimensional (1D) patterns were obtained, as shown in Figure 11(c). Two complete light patterns could be obtained on a sensor in most cases. However, when the finger was close, one of the two images might exceed the sensor. As mentioned before, the strategy of allowing some of the images to exceed the sensor can help to avoid excessively long sensor for compact system volume. Anyhow, more than two images had been obtained to detect finger position.

According to the detection algorithm introduced in Section 3, central points of complete light patterns and dark patterns were found and marked red in Figure 11(c). Chief rays were drawn with the method illustrated in Figure 9, then finger

positions were achieved via average positions of intersection points produced by pair-wise lines. Corresponding to the six cases, finger positions were listed below: ①(x=-0.22cm, y=0.06cm); ②(x=-0.25cm, y=5.07cm); ③(x=-0.20cm, y=10.16cm); ④(x=5.18cm, y=-0.09cm); ⑤(x=5.20cm, y=5.11cm); ⑥(x=5.31cm, y=9.92cm). Figure 13 shows detected and actual finger positions comparatively. Standard deviation of detective errors is 0.41mm. For resolution of 25ppi, each point corresponds to 1.02mm, hence standard deviation of 0.41mm corresponds to only 0.40 points. Therefore, relative detective error cannot cause error that users on the other side of human-computer interface can perceive.

To verify resolution performance, which is another key performance of virtual mouse, the finger was moved along x (lateral) and y (vertical) axis direction for 1cm respectively under the six cases. For instance, corresponding to case ①, the finger was firstly moved from (x=-0.5cm, y=0cm) to (x=0.5cm, y=0cm), and then from (x=0cm, y=-0.5cm) to (x=0cm, y=0.5cm). Before and after every moving operation, positions of central points of all the complete light patterns and dark patterns were recorded. Central point with the largest variation during the finger moving was selected to calculate resolution. Pixel number corresponding to the largest central point variation divided by 1cm (0.39inch) is just the value of resolution. Table 1 shows experimental results and the minimum resolution over the whole working area was 13ppi. It must be pointed that pixel size of 4μm was actually used in the experiments as we were not able to get higher class sensor. If sensor with 2μm pixel size is used, the minimum resolution will be naturally doubled to 26ppi, which is higher than that of conventional touchpad or touchscreen (25ppi).

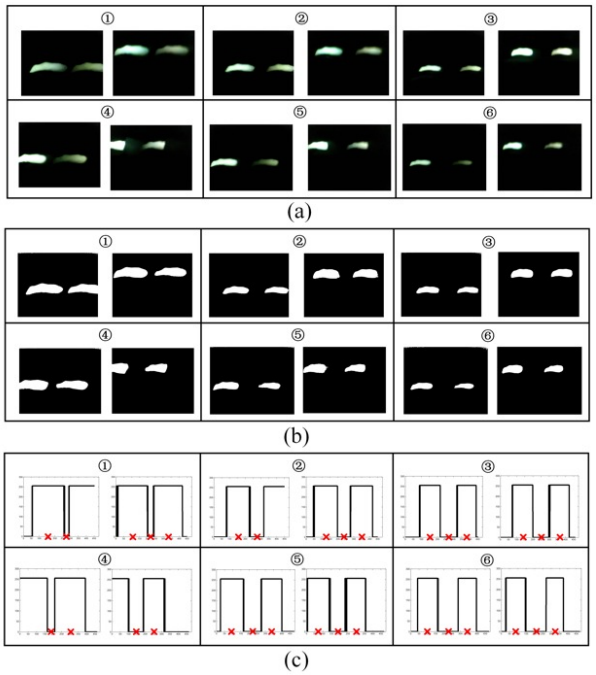

Figure 11. Corresponding to six finger positions: (a) Raw signals; (b) Binary images after preliminary process; (c) Extracted 1D patterns where red marks denote central points of complete light patterns and dark patterns

Table 1. Measured lateral and vertical resolution around six representative finger points in the working area

Finger position	Lateral		Vertical	
	Pixel number	Resolution	Pixel number	Resolution
①	50	127ppi	10	26ppi
②	30	76ppi	7	18ppi
③	15	38ppi	5*	13ppi*
④	52	133ppi	14	36ppi
⑤	31	79ppi	10	26ppi
⑥	15	38ppi	7	18ppi

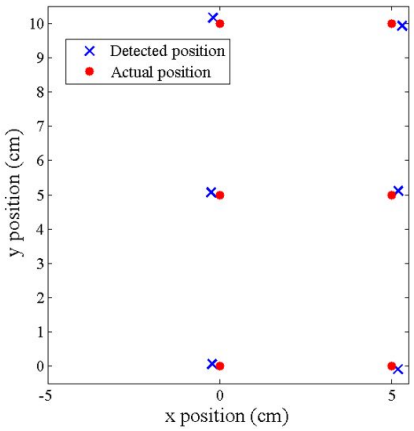

Figure 12. Detected and actual finger positions in working area

5. DISCUSSION AND FUTURE WORK

Proposed virtual mouse prototype with dimension of 3.1mm (thickness) × 4.5mm (length) × 2 has been verified to have a minimum resolution of 25ppi if sensor with 2μm pixel size is used. Then it is capable to replace current touchpad or touchscreen. Further, according to our resolution calculation, if system thickness is relaxed to 25mm and 2μm pixel size is still used, resolution higher than 200ppi can lead our virtual mouse to replace real opto-mechanical mouse. If system is required to be smaller further, sensor with smaller pixel size, like 1μm, can be adopted. In that case, resolution will be increased proportionally with respect to pixel size. Then the gap can be narrower to trade-off between resolution and system volume

Additionally, there is no technical barrier while converting this 2D virtual mouse into a 3D one. As long as a planar light sensor is used instead of linear sensor, finger's moving in z-axis direction will be reflected by image's moving in z-axis direction on the planar sensor. Since the situation in x-axis and z-axis direction is identical, resolution along z-axis direction will be equal to that along x-axis direction. Nevertheless, sensor length in z-axis direction should be determined according to the depth requirement of working area. Larger depth calls for longer sensor in z-axis direction to guarantee complete captured images. In fact, patterns in Figure 12(a) and 12(b) could move vertically when the finger was moved along z-axis direction in experiments because we had already adopted a planar sensor. In future work, currently used experimental architecture will be extended to implement 3D virtual mouse. Accuracy and resolution will be investigated for 3D situation. Moreover, gesture control will be discussed as 3D positional information can be captured[17-19].

6. CONCLUSION

In this paper, an optical architecture comprising IR source, two groups of lens arrays and light sensors was proposed to implement virtual mouse. In this architecture, finger position was calculated from images produced by lenses.

For given working area size of 10cm by 10cm and working distance of 12cm, dimension of each group of the device was only 3.1mm (thickness) by 4.5mm (length), which is much smaller than that of camera-based device. Experiments were implemented by using sensors with 4μm pixel size. In combination with proposed detection algorithm with low computations, standard deviation of detective error calculated from six representative finger positions was 0.41mm, which corresponded to only 0.40 distinguishable points that cannot be perceived by users. Minimum resolution over the working area measured in the experiments was 13ppi. If sensor with 2μm pixel size is used, minimum resolution will be doubled to 26ppi, which is higher than that of conventional touchpad or touchscreen. Sufficient accuracy and resolution verified by experiments, along with compact volume, low-computation detection algorithm, and potential anti-interference feature of IR source, demonstrate that proposed virtual mouse can practicably replace current touchpad or touchscreen in purpose of more convenient operation. If system thickness is relaxed to 25mm and 2μm pixel size is still used, resolution higher than 200ppi in the working area can realize replacing real opto-mechanical mouse.

ACKNOWLEDGEMENTS

This work was supported by Coretronic Corp., Taiwan and Ministry of Science and Technology (MOST) of R.O.C. projects under grant number 102-2221-E-009-167-MY3 and 102-2221-E-009-168-MY3.

REFERENCES

[1] Xu, J., and Zhang, X., "Finger mouse system based on computer vision in complex backgrounds," Proc. SPIE 9067, 906707-1-906707-5 (2013).
[2] Bhowmik, A. K., "Advances in interactive display technologies," J. Soc. Inf. Disp. 20(8), 409-412 (2012).
[3] Lee, S. C., Li, B. H., and Starner, T., "AirTouch: Synchronizing in-air hand gesture and on-body tactile feedback to augment mobile gesture interaction," 15th Annual Int. Symp. on Wearable Computers (ISWC), 3-10 (2011).
[4] Ortega, M., and Nigay, L., "AirMouse: finger gesture for 2D and 3D interaction," Human-Computer Interaction - INTERACT 5727, 214-227 (2009).
[5] Lee, K., Kim, Y. S., Wang, Y., and Kim, Y. S., "A remote pointing device by using modulated IR signals with a 2-D striped pattern," IEEE Transactions on Consumer Electronics 59(3), 699-704 (2013).
[6] Fan, H. Q., Du, Y. H., and Su, Z. T., "InfTouch: A multi-touch and free-air interaction technique for over-sized displays," 6th International Congress on Image and Signal Processing (CISP) 3, 1452-1456 (2013).
[7] Liou, J.-C., and Hsu, C.-C., "Floating display with interactive virtual touch module," SID Int. Symp. Digest Tech. Papers 44, 19-20 (2013).
[8] Kent, J., "Touch-technology diversity in commercial applications," SID Int. Symp. Digest Tech. Papers 45, 615-618, (2014).
[9] Wilson, A. D., "Using a depth camera as a touch sensor," Proc. ACM International Conference on Interactive Tabletops and Surfaces, 69-72 (2010).
[10] Lin, S.-Y., Shie, C.-K., Chen, S.-C., and Hung, Y.-P., "AirTouch panel: a re-anchorable virtual touch panel," Proc. 21st ACM International Conference on Multimedia, 625-628 (2013).
[11] Morrison, G. D., "A camera-based input device for large interactive displays," Computer Graphics and Applications 25 (4), 52-57 (2005).
[12] Wang, G.-Z., Huang, Y.-P., Chang, T.-S., and Chen, T.-H., "Bare finger 3D air-touch system using an embedded optical sensor array for mobile displays," J. Disp. Tech. 10(1), 13-18 (2014).
[13] Fang, H.-H., Wang, G.-Z., Chang, C.-W., and Huang, Y.-P., "3D multi-touch system by using coded optical barrier on embedded photo-sensors," SID Int. Symp. Digest Tech. Papers 44, 1513-1516 (2013).
[14] Wang, G.-Z., Huang, Y.-P., and Chang, T.-S., "Bare finger 3D air-touch system with embedded multiwavelength optical sensor arrays for mobile 3D displays," J. Soc. Inf. Disp. 21(9), 381-388 (2013).
[15] Huang, Y.-P., Wang, G.-Z., Chang, T.-S., and Chen, T.-H., "Three-dimensional virtual touch display system for multi-user applications," J. Disp. Tech. 9(11), 921-928(2013).
[16] Smith, W. J., [Modern optical engineering (4th Ed)], McGraw-Hill Education (2007).
[17] Suarez, J., "Hand gesture recognition with depth images: a review," Proc. 2012 IEEE RO-MAN, 411-417 (2012).
[18] Hassanfiroozi, A., Huang, Y.-P., Javidi, B., and Shieh, H.-P. D., "Hexagonal liquid crystal lens array for 3D endoscopy," Opt. Exp. 23(2), 971-981(2015).
[19] Huang, Y.-P., Jena, T.-H., Chang, Y.-C., Shieh, P.-Y., Chen, C.-W., and Liao, L.-Y., "Individually adapted LC-lens array for 3D applications," Molecular Crystals and Liquid Crystals 605(1), 267-274 (2014).