A motion gesture sensor using photodiodes with limited field-of-view

Yong Sin Kim and Kwang-Hyun Baek[*]

School of Electrical and Electronics Engineering, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, 156-756, Korea
*kbaek@cau.ac.kr

Abstract: This paper presents a low-power and small-size motion gesture sensor (MGS) based on an active infrared (IR) proximity sensor. While conventional proximity-based MGSs require two LEDs (the most power-hungry components) and one photodiode, the proposed MGS eliminates one of LEDs at the expense of an additional photodiode as a separate channel. In conjunction with an optical block that limits the field-of-view of each photodiode, the power consumption and area of the proposed MGS can be reduced by up to 52% and 69%, respectively, compared with conventional proximity-based MGSs. Optical simulation and test results validate theoretical analysis presented.

©2013 Optical Society of America

OCIS codes: (130.0250) Optoelectronics; (130.3120) Integrated optics devices; (130.6010) Sensors; (040.5160) Photodetectors; (230.5170) Photodiodes.

References and links

1. S. Mitra and T. Acharya, "Gesture recognition: a survey," IEEE Trans. Syst. Man Cybern. C **37**(3), 311–324 (2007).
2. J. O. Wobbrock, A. D. Wilson, and Y. Li, "Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes," in *Proceedings of ACM Symposium on User Interface Software and Technology* (ACM, New York, 2007), pp. 159–168.
3. H. Lee and J. Park, "Touch play pool: touch gesture interaction for mobile multifunction devices," in *Proceedings of IEEE Conference on Consumer Electronics* (Institute of Electrical and Electronics Engineers, 2012), pp. 291–292.
4. I. Yang and O. Kwon, "A touch controller using differential sensing method for on-cell capacitive touch screen panel systems," IEEE Trans. Consum. Electron. **57**(3), 1027–1032 (2011).
5. J.-H. Kim, N. D. Thang, and T.-S. Kim, "3-D hand motion tracking and gesture recognition using a data glove," in *Proceedings of IEEE International Symposium on Industrial Electronics* (Institute of Electrical and Electronics Engineers, 2009), pp. 1013–1018.
6. Y. Han, "A low-cost visual motion data glove as an input device to interpret human hand gestures," IEEE Trans. Consum. Electron. **56**(2), 501–509 (2010).
7. J. Kela, P. Korpipää, J. Mäntyjärvi, S. Kallio, G. Savino, L. Jozzo, and D. Marca, "Accelerometer-based gesture control for a design environment," Pers. Ubiquitous Comput. **10**(5), 285–299 (2006).
8. E. Y. Ahn, J. H. Lee, T. Mullen, and J. Yen, "Dynamic vision sensor camera based bare hand gesture Recognition," in *Proceedings of IEEE International Symposium on Computer Intelligence for Multimedia* (Institute of Electrical and Electronics Engineers, 2011), pp. 52–59.
9. T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," Comput. Vis. Image Underst. **81**(3), 231–268 (2001).
10. H. Cheng, A. M. Chen, A. Razdan, and E. Buller, "Contactless gesture recognition system using proximity sensors," in *Proceedings of IEEE Conference on Consumer Electronics* (Institute of Electrical and Electronics Engineers, 2011), pp. 149–150.
11. A. M. Chen, H. T. Cheng, A. Razdan, and E. B. Buller, "Methods and apparatus for contactless gesture recognition," Qualcomm Inc., U.S. Patent App. 13/161,955 (2011).
12. M. Igaki, H. Osawa, and T. Tsuchikawa, "Illumination device," Rohm Co. LTD., U.S. Patent App. 13/187,593 (2011).
13. T. Chang, K. P. Wu, C. J. Fang, C. T. Chan, C. T. Chuang, and F. Y. Liu, "Light sensor system for object detection and gesture recognition, and object detection method," U.S. Patent App. 13/494,000 (2012).
14. Silicon Labs white paper, "Infrared gesture sensing" (Silicon Laboratories Inc., 2011), http://www.silabs.com/Support%20Documents/TechnicalDocs/AN580.pdf
15. T. Martinez, D. Wick, and S. Restaino, "Foveated, wide field-of-view imaging system using a liquid crystal spatial light modulator," Opt. Express **8**(10), 555–560 (2001).
16. R. Winston, J. C. Miano, and P. G. Benitez, *Nonimaging Optics* (Academic, 2004).

17. I. Luque-Heredia, J. Moreno, P. Magalhaes, R. Cervantes, G. Quemere, and O. Laurent, *Concentrator Photovoltaics* (Springer, 2007), Chap. 11.

1. Introduction

As the daily engagement of portable devices such as smartphones and tablets becomes tremendously increased, various human-machine interfaces (HMIs) are demanded in markets [1]. One of the hot topics in current HMIs of these devices is human motion detection, which incorporates the human factors to interact with computers intuitively. It can be implemented by using motion gesture sensors (MGSs), which can be categorized into four types: touch-based, motion-based, vision-based, and proximity-based systems. The touch-based system is one of the most natural and intuitive technology either using fingers or a pen [2–4]. However, when wearing gloves or getting hands wet or dirty, sensing by touch may not be performed. A motion-based system supports single-handed interaction by having an external controller at the expense of an additional cost [5,6]. A gyroscope or accelerometer can also be used for a motion-based system, but it causes a shaky screen [7]. The vision-based system using an embedded camera allows users to make basic gestures for interfacing without touching the device but requires high-power consumption [8,9]. Recently, proximity-based MGSs have emerged in the market because of their intuitive control methods for detecting physical movement of an object without touching the device. In order to be incorporated with portable devices, low power and small size are two key features that lead an active IR proximity-based MGS to be the most profitable scheme [10].

Conventional active IR proximity-based MGS systems are composed of at least two LEDs and a photodiode (PD) for conducting phase-based sensing [10–14]. When IR light is emitted from each LED alternately, it bounces off an object and reaches onto a photodiode. The two main drawbacks of this method are: using multiple LEDs that dissipate large amount of power and requiring large space between LEDs for achieving a certain recognition rate. In this paper, we propose an active IR proximity-based MGS using a single LED. It reduces power consumption with an additional photodiode as a separate channel. In order to implement all components close to each other for minimizing hardware formation, the proposed scheme uses a small optical block that limits the field-of-view (FOV) of each photodiode.
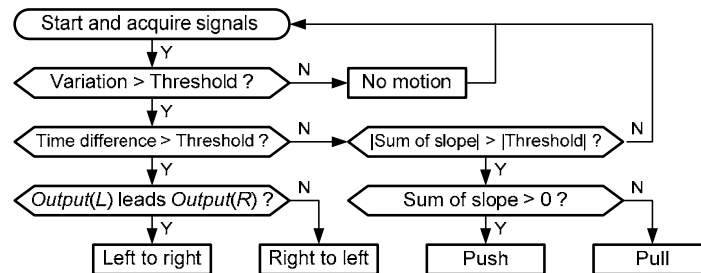

Fig. 1. The flowchart of a conventional MGS for detecting four different gestures: Left to right, right to left, push, and pull.

2. System overview

Conventional proximity-based MGS systems determine the approximated position of a reflective object repeatedly by using simple optical components such as multiple IR LEDs and a photodiode. Figure 1 shows the flowchart of an MGS algorithm to recognize some distinct hand motions such as left to right, right to left, push, and pull. When the process starts, two separate signals are acquired for single-axis motion detection. If the variation of these signals is above a certain threshold level, the timing information of two signals is obtained to check which one leads the other. If the signal *Output*(L) generated from the left-hand side LED leads the signal *Output*(R) from the right-hand side LED, the MGS algorithm decides that an object is moving from left to right, and vice versa. Similarly, "push" and "pull" can be determined by acquiring the slope of the signals. One of the key properties that vary the performance of an

MGS system is the FOV of each optical component. Here we present the theoretical analysis for two different types of MGS systems.
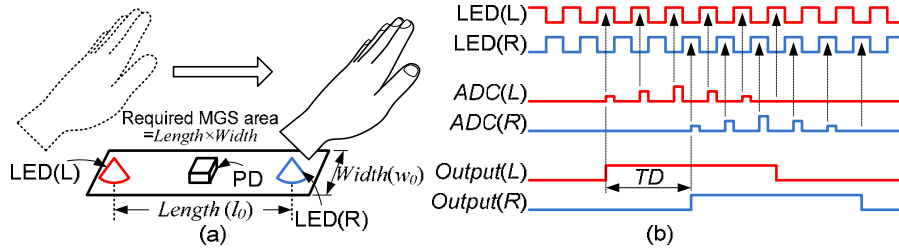

Fig. 2. Conceptual diagram of a conventional active IR proximity-based MGS using a phase-based sensing algorithm for detecting the motion of a hand. (a) Hardware formation. (b) Output timing diagram for detecting a hand moving from left to right.

2.1 An MGS using a phase-based sensing algorithm

The basic concept of a conventional proximity-based MGS system using a phase-based sensing algorithm is shown in Fig. 2(a). In order for a photodiode to differentiate signals from two LEDs, its output is divided into two phases that lead each LED to be turned on and off in succession. The intensity of reflected IR light changes according to the distance and the angle between a hand and LEDs. Figure 2(b) shows timing diagrams of two separate phase signals when a hand moves from left to right. Two LEDs turn on and off alternately at a given modulation frequency. Then the two phase outputs of the photodiode are amplified, filtered, and converted into a digital signal by an analog-to-digital converter (ADC). The output signals, *ADC*(*L*) and *ADC*(*R*), need to be return-to-zero format in order to be ready for the other phase. Finally, *Output*(*L*) and *Output*(*R*) are determined to obtain the edge of each phase signal. In this case, *Output*(*L*) leads *Output*(*R*), indicating that a hand moves from left to right. By counting the number of pulses between the rising edges of two output signals, we can obtain the speed of the movement. Note that timing difference *TD* is proportional to the distance between two LEDs, which requires the minimum distance of l_0 [14]. Thus, in order to have a higher recognition rate, a larger area for hardware formation is required.

Figure 3(a) depicts the conical emission of each LED in a conventional proximity-based MGS. Since each emission is symmetrical across the vertical line at the center of each component, the FOVs of LEDs and the photodiode are equal to the emission angle α and the view angle β, respectively. The overlapped area of the emission angles of two LED is defined as the ambiguous zone, where the photodiode receives the reflected signals from both LEDs simultaneously. If an object moves within the ambiguous zone, the motion detection algorithm fails to make a decision. A cross-sectional view of the conical emission is depicted in Fig. 3(b). In the figure, the detectable zone (L) and (R) represent the zone where an object reflects IR light that comes only from left or right LEDs, respectively. Each gesture can be detected by measuring the timing difference between the entry into detectable zone (L) and (R). The length *l* of a detectable zone at the height *h* can be defined as

$$l = l_0, \text{ where } h \geq h_0 = \frac{0.5l_0}{\tan(\alpha/2)}. \tag{1}$$

It indicates that the length of a detectable zone is the same as the distance between two LEDs. Assuming that a hand moves at a constant speed, *v*, *TD* can be expressed as

$$TD = l / v = l_0 / v, \tag{2}$$

where *TD* is constant regardless of the values of α and β if $\alpha \leq \beta$. As *h* increases, however, less amount of reflected light proportional to $1/h^2$ reaches on the photodiode, which may lead to a lower recognition rate.
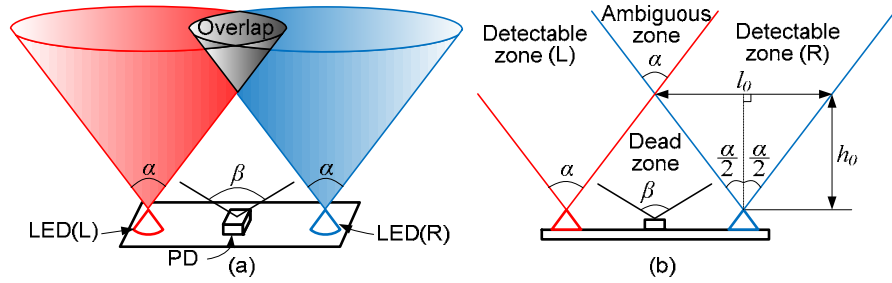
Fig. 3. Various zones according to the emission and acceptance angles of each component in a conventional active IR proximity-based MGS shown in Fig. 1. (a) Conical emission of each LED. (b) FOV of each component, where $\alpha \leq \beta$.

2.2 The proposed MGS using a channel-based sensing algorithm

The proposed MGS is developed in two stages. Stage one is low-power development of an MGS by using a channel-based sensing algorithm. It consists of a single LED and two photodiodes as shown in Fig. 4(a). When the LED is on, two photodiodes independently receive reflected light from a hand and form two separate channels that measure the amount of reflection. If α is set to be smaller than β as described in Fig. 3, the detectable zones becomes negligible since the LED emits IR light only within the ambiguous zone. In order to keep this from happening, the relationship between α and β in Eqs. (1) and (2) should be the other way around as

$$l = l_0, \text{ where } h \geq h_0 = \frac{0.5l_0}{\tan(\beta/2)} \text{ and } \alpha \geq \beta. \tag{3}$$

Figure 4(b) depicts the timing diagram for the MGS using a channel-based sensing algorithm. Since having only one LED, the signals from two separate photodiodes are amplified through two separate channels and converted into digital signals by using their own ADCs. Thus, the outputs of ADCs do not need to be return-to-zero format. Other functions are identical to the conventional scheme.
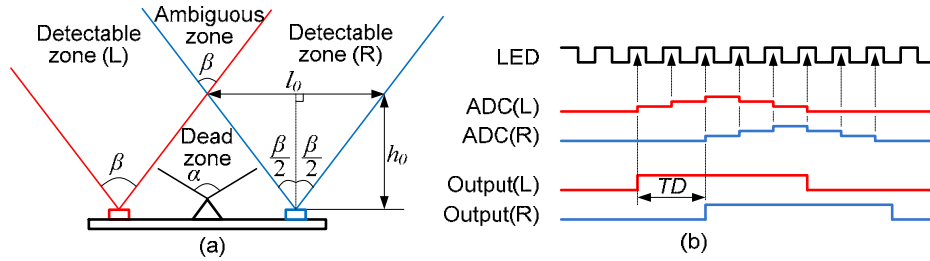

Fig. 4. The first stage development of the proposed MGS using a channel-based sensing algorithm. (a) FOV of each component, where $\alpha \geq \beta$. (b) Timing diagram when a hand moves from left to right.

The second stage of development is reducing the area of the first stage MGS by putting all the components next to each other. However, TD in Eq. (2) becomes zero by simply doing so because of the very short distance between two photodiodes. To overcome this problem, we propose a technique that limits FOVs of each photodiode. In a conventional imaging system, complex designs utilizing multiple optics may be required for controlling FOV [15]. However, the proposed MGS is a nonimaging system that considers rays of light as the path along which light energy travels with surfaces reflecting or blocking the light [16]. In our case, the FOV of each photodiode can be controlled without using complicated optics. As an example, Fig. 5(a) illustrates the simplified diagram of a sun tracking sensor that consists of

two photodiodes and an optical block for single axis tracking [17]. When placing an optical block between adjacent photodiodes for blocking rays of light, different photocurrents are generated depending on the incident angle of lights. Then the difference allows a tracker to move toward the direction of higher valued one until they are balanced. Applying this concept to the proposed MGS, all the components can be put close together as shown in Fig. 5(b). The detailed 2D cross-sectional view is depicted in Fig. 5(c), where γ is the half angle of FOV limited by the optical block. Assuming that the distance between two photodiodes and the height of the optical block are negligible compared with h, l can be approximated as shown in Fig. 5(d):

$$l = h(\tan(\beta/2) - \tan\gamma), \tag{4}$$

where the emission angle α of the LED needs to be larger than β. Unlike conventional MGSs, there are two additional advantages of the proposed scheme. First, TD increases as h does since the angle of the overlapped zone is smaller than the view angle of a photodiode ($2\gamma<\beta$). Second, the height of the dead zone is very small compared with the conventional scheme, which releases one design constraint of an MGS system.
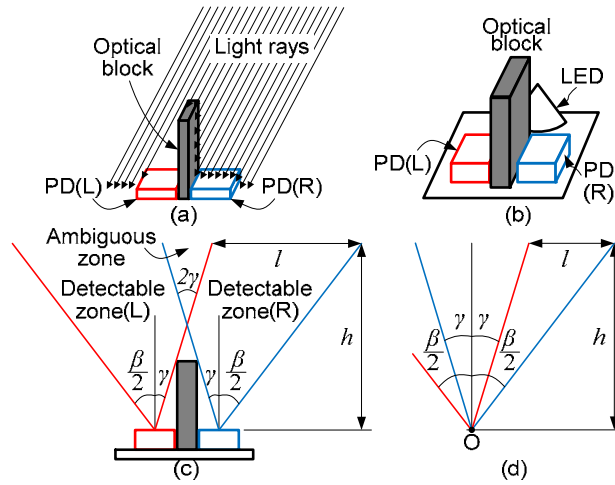

Fig. 5. The second stage development of the proposed MGS using a channel base sensing algorithm with an optical block. (a) An example of an optical block used for a sun tracking sensor with a pair of photodiodes. (b) Component formation of the proposed scheme. (c) 2D cross-sectional view. (d) Approximated FOVs for calculating the length of detectable zone.

3. Simulation results

An MGS discussed in Figs. 2 and 3 inherently suffers from a large overlapped zone that may lead to functional error when an object moves within the zone. Moreover, sweeping a hand or a finger across the dead zone may cause a lower recognition rate even though the timing difference is high enough. Figure 6(a) shows the simulation results for the length of detectable zone according to three different values of β. The distance l_0 between photodiodes is set to 80 and 20 mm for a hand and finger, respectively. The height of the dead zone, h_0, becomes smaller as β gets large. On the other hand, the dead zone caused by the proposed scheme is almost negligible as shown in Fig. 6(b). In addition, the area required for conventional MGS systems varies by the width of an object [14]. However, the proposed system can be incorporated with the various sizes of an object, which makes the system universal.

For optical simulations, *LightTools* software is used to process ray tracing and to predict luminance distributions. To quantify the amount of light received onto each photodiode depending upon the location of an object, the object was assumed to be a light source without using an LED. The geometrical dimensions of the simulation environments are shown in Fig.

7(a). An object sized 80 mm by 80 mm is located 100 mm away from the sensor. One million non-sequential rays are emitted at a random angle within β, which allows the FOV of each photodiode to be confined by the object. Since the ratio between the half-width of a photodiode and the height of the optical block is 1 to 7.5, γ is set to be $\gamma = \tan^{-1}(0.7/5.3) \approx 7.5°$. As the optical block becomes taller, the length of the detectable zone increases. However, the increased height of the overall system may not be suitable for some portable applications. The luminance distributions ($\beta = 60°$) are investigated while changing the position of the object horizontally as depicted in Fig. 7(b). Each graph is normalized to its own maximum value. The length of each detectable zone can be extracted at the half-maximum points, $L_1 = 42.8$ mm and $L_2 = 39.9$ mm, which shows in good accordance with the simulated detectable length marked as point **A** in Fig. 6(b).
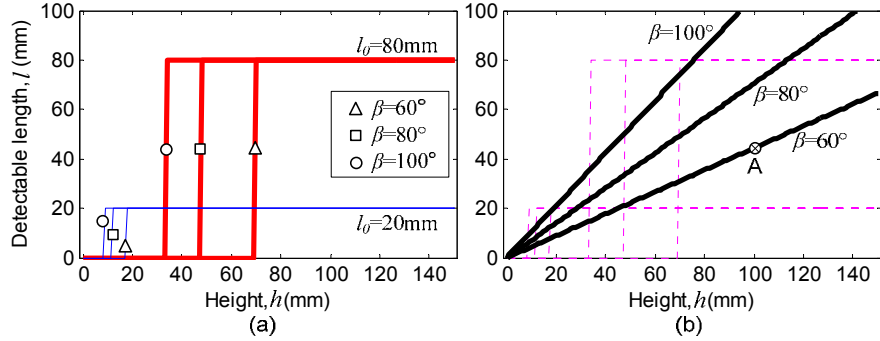

Fig. 6. Simulation results for the length of detectable zones according to three different values of β. (a) Conventional MGSs with a hand ($l_0 = 80$ mm) and a finger ($l_0 = 20$mm). (b) The proposed scheme with an optical block at $\gamma = 7.5°$. Dashed lines are for conventional schemes from Fig. 6(a).
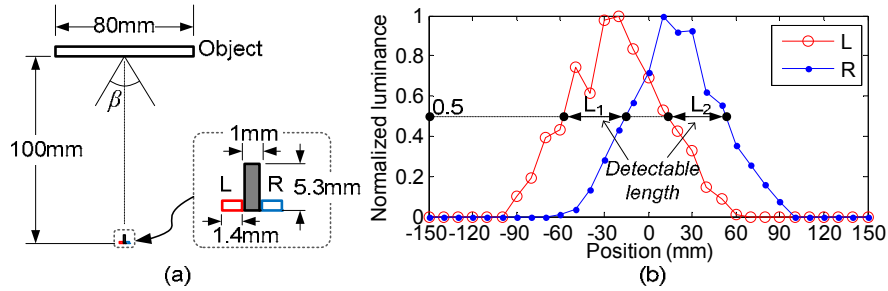

Fig. 7. Optical simulation using *LightTool*s. (a) Simulation setup. (b) Luminance distributions with $\beta = 60°$ and $\gamma = 7.5°$.

4. Experimental results

Figure 8(a) depicts a block diagram of the proposed MGS. As the reflected IR light generates photocurrent in each photodiode, a transimpedance amplifier (TIA) converts the current to the form of voltage. Then the signal is amplified and filtered to reduce the effects of ground bouncing and power noise. For prototype testing, an optical block was fabricated by using 3D printing and painted black to prevent any light path through. When an LED, sized 1.6mm by 3.2mm, is adopted by using a surface-mounted package ('type no. 1206'), the area of the conventional MGS system that detects the motion of an object can be calculated as 1.6 mm by 80 mm [14]. On the other hand, the area of the proposed MGS can be reduced to 4 mm by 10 mm as shown in Fig. 8(b). Thus, the area reduction ratio of the proposed scheme becomes (1.6 × 80–4 × 10)/(1.6 × 80) = 0.69. The power consumption P_{prop} of the proposed MGS is

dominated by the IR LED compared with other balance-of-system (BOS), which can be defined as

$$P_{prop} = P_{LED} + P_{BOS} = (D \cdot I_{LED} + \bar{I}) \cdot V_{DD}, \tag{5}$$

where P_{BOS} is power consumption in BOS, D is a duty ratio, I_{LED} is the peak current through the LED, $\bar{I}$ is the average current of BOS, and V_{DD} is the supply voltage. When the LED is in active mode, it consumes average power of 12.5 mA at $D = 0.5$, $I_{LED} = 25$ mA, and 2 kHz modulation frequency. The average current of BOS is controlled under 1mA. Assuming that P_{BOS} is identical for both proposed and conventional MGS systems, the power reduction ratio PR of the proposed scheme becomes

$$PR = \frac{P_{conv} - P_{prop}}{P_{conv}} = \frac{D \cdot I_{LED}}{D \cdot I_{LED} \cdot 2 - \bar{I}} \leq \frac{0.5 \cdot 25 (mA)}{0.5 \cdot 25 (mA) \cdot 2 - 1 (mA)} = 0.52, \tag{6}$$

where P_{conv} is power consumption in the conventional MGS system.

Figures 8(c) and 8(d) show analog outputs of the proposed MGS before and after placing the optical block. Timing difference was measured at the desired moving speed of an object by using a white card sized 80 mm by 80 mm. The difference in analog outputs without the optical block is almost negligible since the time difference is proportional to the distance between two photodiodes. After placing the optical block between two photodiodes ($\beta = 60°$), the timing difference can be clearly seen. The average value of timing differences after 20 iterations yields 32.7 ms. Since the timing difference is also the function of the velocity of a moving object, it is hard to quantify the performance of the proposed optical blocks. Thus, we conducted the static positioning tests of an object to quantify the length of detectable zone at a given height, $h = 100$ mm. The test procedures are as follows:

1. Position the object at a certain location.

2. Measure the analog DC outputs of two channels.

3. Move the object to next location.

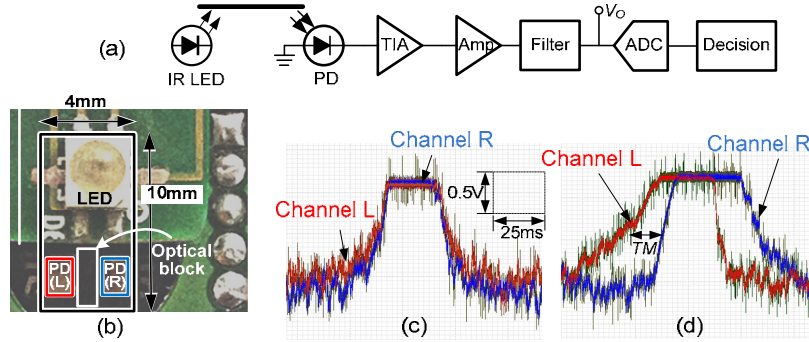4. Go to step 2 until all the points within a desired range are measured.

Fig. 8. Test results for the proposed MGS. (a) Simplified block diagram of the proposed MGS (only single channel is shown). (b) Optical component formation. (c) Analog outputs for swiping from left to right without an optical block. (d) Analog outputs for swiping from left to right with an optical block.

Since this work is focused on limiting the FOV of each photodiode, the detectable length, which is proportional to the recognition rate, is measured with different heights (related to γ) of the optical block at given α and β ($\alpha = \beta = 60°$). Note that $\alpha \geq \beta$ and $\beta \geq 2\gamma$ from Eq. (1) to Eq. (4) for the proposed MGS. Figure 9 shows the lengths of detectable zone, which are defined as L_1 (left-hand side) and L_2 (right-hand side) at three different γ. The values of L_1 and

L_2 are (0,0), (9.1 mm, 10.0 mm), and (22.6 mm, 22.7 mm) for $\gamma = 30°$, 18.2°, and 7.5°, respectively. As indicated in Fig. 9, the FOV of each photodiode gets more limited as the optical block becomes taller, which increases the length of detectable zone as well as the recognition rate. It should be also noted that the measured results are different from the simulation ones shown in Fig. 6 and Fig. 7, because the intensity of reflected light according to the emission angle is not considered in the simulations.

The recognition accuracy of the proposed system is evaluated by sweeping two white-colored cards which mimic the human hand and finger. The widths (w) of two cards are set to 80 mm (hand) and 2 0mm (finger). In order to also investigate the dependency of the recognition rate upon the distance (h) between an object and the sensor, three h values of 50 mm, 75 mm, and 100 mm are used. Therefore, the total 6 sets of test conditions with different w and h values are performed, and each set is repeatedly measured 500 times to obtain an average value of the recognition rate. Table 1 shows the minimum sampling frequency to achieve the recognition rate of 99.5%. Motion gestures are not able to be detected when $w = 20$ mm and $h \geq 75$ mm as shown in Table 1, because narrower objects cause less amount of IR light reflected from the objects. It would decrease the signal-to-noise ratio (SNR) as h increases. For comparison, the conventional MGS in [10] achieved recognition rate of 98%. However, the dependencies of the recognition rate upon w and h are not quantified.
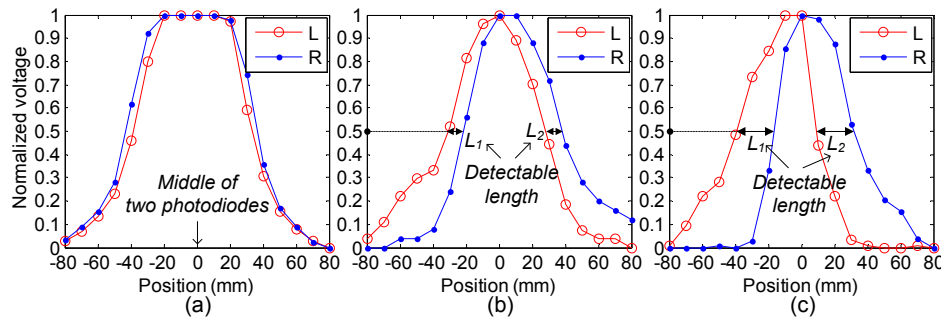

Fig. 9. Outputs of the proposed MGS as a function of object's position with three different values of γ. (a) $\gamma = \beta/2 = 30°$. (b) $\gamma = 18.2°$. (c) $\gamma = 7.5°$.

Table 1. Minimum Sampling Frequency Required to Achieve the Recognition Rate of 99.5%

h (mm) \ w (mm)	25	50	75	100
20	0.1 kHz	0.11 kHz	0.24 kHz	1.54 kHz
80	0.18 kHz	0.29 kHz	N/A	N/A

5. Conclusion

Since the detection timing margins in conventional proximity-based motion gesture sensors are proportional to the distance between two LEDs, high power consumption and large footprint cannot be avoided in their design. Each LED consumes several tens of times more power than other balance-of-system in general. In order to reduce power consumption, a novel proximity-based motion gesture sensor using a single LED is proposed at the expense of additional photodiode as a separate channel. The proposed scheme uses a small optical block made out of cheap plastic material for limiting its field-of-view and thereby placing all the components next each other. By using the proposed scheme, the power consumption and the size can be reduced by up to 52% and 69%, respectively, compared with a conventional proximity-based MGS system using two LEDs.

Acknowledgments