Hand Gesture Recognition
with Inertial Sensors and a Magnetometer

Mi-So Lee, Kyung-Won Kim, Mun-Ho Ryu,[1,2*] and Je-Nam Kim[3]

Department of Healthcare Engineering, Chonbuk National University,
567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do 561-756, Republic of Korea
[1]Division of Biomedical Engineering, Chonbuk National University,
567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do 561-756, Republic of Korea
[2]Research Center of Healthcare & Welfare Instrument for the Aged, Chonbuk National University,
567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do 561-756 Republic of Korea
[3]R&D Division, Chonbuk National University Automobile-parts & Mold Technology Innovation Center,
67 Yu-Sang-ro, Deokjin-gu, Jeonju-si, Jeollabuk-do 561-844, Republic of Korea

Due to their low cost and compact size, micro-electromechanical system-based inertial sensors have been utilized to detect human motion. This study presents a hand gesture recognition method with a three-axis accelerometer, gyroscope, and magnetometer attached to a wrist. First, sensor signals were applied to calculate the sensor's orientation with a Kalman filter. These data were used to convert the sensor-frame acceleration signal to a global-frame acceleration signal (orientation calculation). Acceleration of motion was calculated by subtracting gravity from the global acceleration signal. The start and end of hand motion was detected by motion acceleration (motion segmentation). Segmented hand motion was recognized by gesture (gesture recognition). Six gestures (up, down, left, right, click, and double-click) were selected for implementation of the system, and the performance was evaluated.

1. Introduction

Recently, interest in human–computer interaction (HCI) has increased. Consequently, systems to identify user intentions have been actively studied within the proactive computing field, where information is provided actively. In particular, among such studies, a focus has been on gesture recognition applications.[1] Studies on gesture-based user interfaces have been carried out globally to support a natural and comfortable interface between users and various devices.[2] In the gesture recognition field, a wide variety of expressions available with hand gesture recognition has potentially unlimited power.[3] Gesture recognition is an important part of sign language, and it is the basis of communication, remote medical systems, various electronic devices, interface configuration, and the control of the motion of robots.[4]

There are two methods of gesture recognition: non-contact and contact.[5] The non-contact method ensures natural movements by eliminating unnecessary equipment from the user experience. However, with this method it is difficult to track the features of the user's movement because of the

---

*Corresponding author: e-mail: mhryu@jbnu.ac.kr

influence of light, surrounding objects, distance limitations, and other variables.[6] Cheap products such as Kinect v2 from Microsoft were recently launched. This is very advantageous in terms of cost. However, images obtained from kinect2 are unclear in the visual aspect, because the depth of the images becomes inaccurate over time. Kinect v2 fails in both accuracy and resolution.[7]

On the other hand, a magnetometer and an inertial-sensor-based gesture recognition system uses acceleration and angular velocity. This system is not affected significantly by the environment, and its processing speed is faster than other systems. However, in most studies of gesture recognition using inertial sensors, gestures are made in a plane. This limits three dimensional motional operations, making it difficult to configure the interface to a variety of challenging motions.[8]

This study presents a gesture recognition system not limited to a plane. The system discerns three dimensional orientation in a global coordinate system.[9] Knowing the direction, the system can recognize gestures in any plane or at any angle to a plane regardless of the gesture. This even includes gestures for configuring a mouse, which is a common interface between computer and human. Six gestures (up, down, left, right, click, and double-click) were selected for implementation of the system, and the system's performance was evaluated.

2. Materials and Methods

2.1 Sensor

The device is composed of a two-part sensor and a microcontrol unit (MCU) consisting of an inertial sensor and a magnetometer (Fig. 1). A three-sensor three-axis accelerometer, a three-axis magnetometer, and a three-axis gyroscope (L3GD20, STMicroelectronics) were used to develop the device. For communication, an integrated circuit communicator was implemented with an MCU at 400 kHz. The MCU used (STM32F103C8, STMicroelectronics) transmitted mainly sensor data via a universal synchronous/asynchronous receiver transmitter between the PC and the MCU. Sensor data was transmitted at 100 Hz with a timer interrupt function.

2.2 Host application

All data was converted to the correct format: acceleration units for gravity, Gauss measurements for magnetic field units, and degrees per second for gyroscope units. Next, sensor data were calibrated, as each sensor axis in the sensor device was different from that in host application. Finally, the data were processed by applying a moving average filter whose window size was set to 20 units.
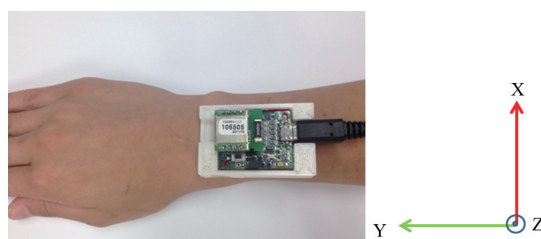

Fig. 1.   (Color online) Sensor device and three-axis coordinate system.

A Kalman filter was used to improve function by sensor fusion; in this case, we conflated accelerometer and gyroscope sensors. Using the trade-off characteristics of the accelerometer and the gyroscope, the sensor mixture compensated for the gyroscope by using the accelerometer if required. Because only the quaternion can be used for Kalman filtering, data from the accelerometer and gyroscope were converted to a quaternion that was a state variable and able to be used in Kalman filtering. Estimations were calculated by subtracting predicted values multiplied by the Kalman gain from measured values. Estimated data were accurate data for position information (Fig. 2).[10,11]

2.3   Gesture recognition

The method for recognizing a gesture is composed of three steps. Acceleration data in the sensor frame was coordinated to data in the global-frame. The system had one viewpoint through this calculation because different viewpoints disappear (Fig. 3).

The first step was segmentation. To separate acquired movement periods from the total data, segmentation used a threshold to detect movement. The threshold was calculated from the norm of normalized acceleration data and was set to 0.15.

The second step was to determine the most significant activation axis because an activated axis is fixed. An activated axis was detected using the gap between minimum and maximum values. If the *Z*-axis was detected, then the system would check the percentage difference between the *Y*-axis and the *Z*-axis. Then, if the percentage difference exceeded 50%, the system would detect that the *Y*-axis was an activated axis; otherwise, the system would detect the *Z*-axis as an activated axis.

Finally, the system used template matching to recognize hand gestures. In the detected movement period, the system found the middle point and then calculated the average between 10% of the period on the basis of the middle point. When the *X*-axis was detected as an activated axis, the positive average was right or left. The same goes for the *Z*-axis: positive was up or down. In the *Y*-axis, the system counted the peak point and then the gesture was separated on the basis of the count number. If the count number was lower than three, then the gesture was recognized as a
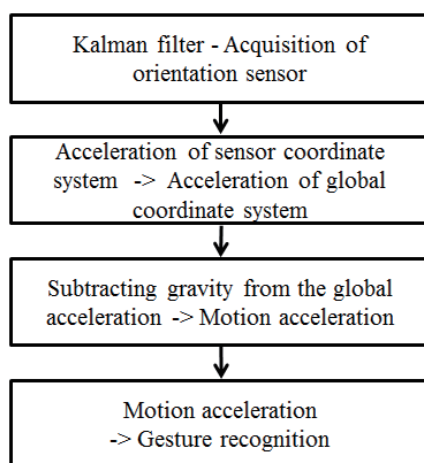

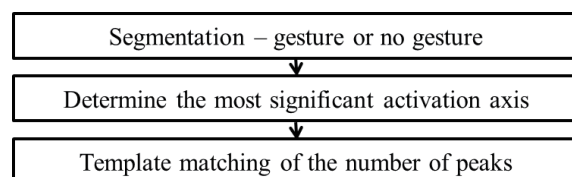Fig. 2.   Host application.

Fig. 3.   Gesture recognition.

click. Alternatively, if the count number was more than three, then the gesture was recognized as a double-click (Fig. 4).

2.4 Experimental protocol

Eight healthy subjects aged 22–24 participated in the experiment. Subjects included four females and four males. An experimental set was composed of five trials of one gesture and had two parts: 90 and 45°. The subjects carried out six gestures, i.e., up, down, left, right, click, and double-click (Fig. 5). The gesture was performed by moving the forearm about 20 cm. Before the experiment, every subject was checked to confirm understanding of the procedure. The period for an experimental set was 10 s, and each motion proceeded for about 1.5 s. Break time between the gestures was about 0.5 s.

3.   Results

The experimental results are shown in Table 1. Six gestures were performed a total of 60 times (90°-5 times, 45°-5 times). Recognition rates for the subjects were 100, 98.33, 100, 100, 100, 95, 100, and 98.33%. For an average recognition rate of 98.75%. In addition, the recognition rate for each gesture was 100, 100, 98.75, 100, 98.75, and 96.25%.

Table 2 shows the results for Subject 6, who showed the largest error. Double-clicking is a gesture of quickly clicking two times. The Subject 6 began the second click before the end of the first click. For this reason, the gesture was not recognized correctly.
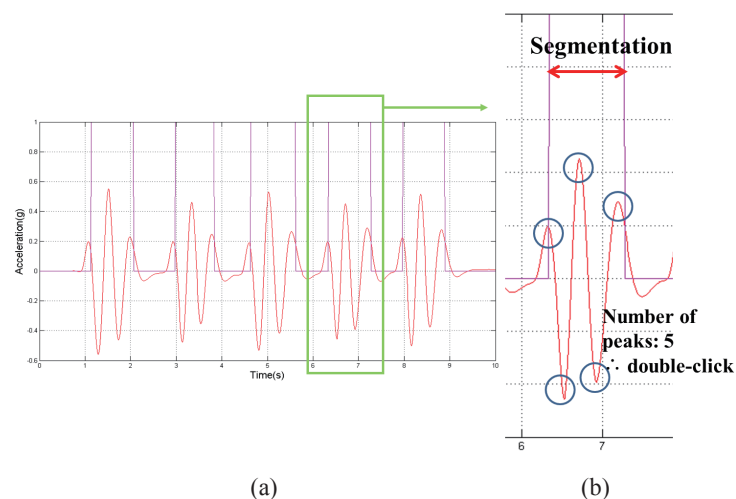

Fig. 4.   (Color online) Acceleration for (a) double-click and (b) gesture recognition.
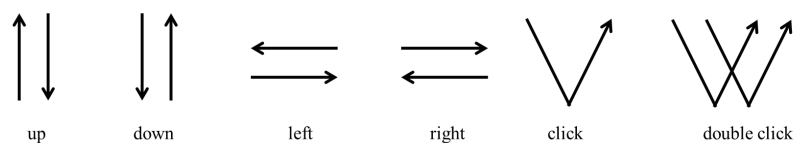

Fig. 5.   Six gestures.

Table 1
All results for recognized gestures (%).

	Up	Down	Left	Right	Click	Double click	Total
Subject 1	100	100	100	100	100	100	100
Subject 2	100	100	100	100	90	100	98.33
Subject 3	100	100	100	100	100	100	100
Subject 4	100	100	100	100	100	100	100
Subject 5	100	100	100	100	100	100	100
Subject 6	100	100	90	100	100	80	95.00
Subject 7	100	100	100	100	100	100	100
Subject 8	100	100	100	100	100	90	98.33
Total	100	100	98.85	100	98.75	96.25	

Table 2
Confusion matrix for Subject 6.

Actual	Recognized					
	Up	Down	Left	Right	Click	Double click
Up	10	0	0	0	0	0
Down	0	10	0	0	0	0
Left	0	0	9	1	0	0
Right	0	0	0	10	0	0
Click	0	0	0	0	10	0
Double click	0	0	0	0	2	8

4.   Conclusions

This report presented a hand gesture recognition system using an inertial sensor. The system consisted of an inertial sensor board and a host application. The sensor board acquired hand acceleration data, and the host application analyzes the data. Six gestures were applied in an experiment to test the system. The test involved eight healthy subjects, and experimental results show a gain in recognition rate from 96.25 to 100%.

The main contribution of this study is the conclusion that there is no limitation on dimension. In the system, sensor-frame data was converted to global-frame data. Therefore, the system uses acceleration data in global-frame coordinates to acquire a unified viewpoint to monitor gesture information. With no limitations on dimension for this system, users are given the convenience of being free to perform any gesture.

Acknowledgements

References

1 S. Mitra and T. Acharya: IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **37** (2007) 311.
2 D. P. Hong and W. T. Woo: Telecommun. Rev. **18** (2008) 403.
3 H. K. Lee and J. H. Kim: IEEE Trans. Pattern Anal. Mach. Intell. **21** (1999) 961.
4 D. Biswas, A. Cranny, A. F. Rahim, N. Gupta, K. Maharatna, N. R. Harris, and S. Ortmann: Biomed. Eng. Lett. **4** (2014) 403.
5 U. X. Tan, K. C. Veluvolu, W. T. Latt, C. Y. Shee, C. N. Riviere, and W. T. Ang: IEEE Sensors J. **8** (2009) 1385.
6 J. H. Kim, N. D. Thang, and T. S. Kim: IEEE Int. Symp. Ind. Electron. **1** (2009) 1013.
7 D. W. Shin and Y. S. Ho: J. Broadcast Eng. **20** (2015) 414.
8 R. Xu, S. Zhou, and W. J. Li: IEEE Sensors J. **12** (2012) 1166.
9 H. Lahamy and D. D. Lichti: Sensors **12** (2012) 14416.
10 J. N. Kim, M. H. Ryu, H. R. Choi, Y. S. Yang, and T. K. Kim: Int. J. Distrib. Sens. Netw. **2013** (2013) 7.
11 D. Roetenberg, H. J. Luinge, C. T. M. Baten, and P. H. Veltink: IEEE Trans. Neural Syst. Rehabil. Eng. **13** (2005) 395.