Hand Gesture Interface using Light-dark changes in an Illuminance meter built in mobile devices

Kento Matsui[1], **Hiroto Aida**[2], **Hikaru Ichikawa**[1], **Mitsunori Miki**[2], **and Hiroki Murakami**[1]
[1]Graduate School of Science and Engineering, Doshisha University, Kyoto, Japan
[2]Department of Science and Engineering, Doshisha University, Kyoto, Japan

Abstract— *This study examines a hand gesture interface using light-dark change in illuminance meters built in smartphones and tablets. A hand gesture is to be detected from a change in illuminance values obtained by an illuminance meter and classified through decision tree learning. This realizes intuitive operations by hand gestures on a mobile device. While many of studies on hand gesture recognition use camera images, it result in large computational complexity required for gesture recognition due to image analysis. Privacy consideration is also required. By using light-dark changes in an illuminance meter, hand gestures can be recognized with less computational complexity and lighter applications can be realized. It also does not require privacy consideration. The result of the precision verification experiment showed that 5 types of hand gestures were recognized at the average recognition rate of 95%. Generic applicability to unknown users and unknown light environments was also confirmed.*

Keywords: hand gesture, natural user interface, illuminance meter

1. Introduction

NUI (natural user interface), which allows intuitive operation, has drawn attention in recent years. NUI is a user interface that allows operation by an intuitive action and realizes interaction with computer using an action such as touch or voice operation. Extensive research has been made on gesture interface above all, among topics on NUI, and many researchers have proposed various methods for gesture recognition [1][2][3]. There has also been a progress in development and commercialization of a device capable of gesture recognition, represented by Kinect [4] and Leap Motion [5]. Technology of gesture recognition has been positively utilized in a variety of areas including entertainment and medicine [6][7][8].

On the other hand, as mobile devices have rapidly spread in recent years, there has been increasing research using information from sensors of such devices. A diverse variety of sensors, including proximity sensor, accelerometer, and geomagnetic sensor, are embedded in smartphones and tablets. These sensors are used for various research and services including user behavior estimation and indoor location estimation.

There are studies on gesture interface also among research using mobile devices. Above all, there is extensive research on hand gesture interface for the recognition of hand movement. Various recognition methods have been proposed including a method for recognizing desk rubbing gestures by using accelerometers and microphones [10], a method for recognizing hand gestures using an RGB camera of a mobile device [11], etc. These methods, however, have issues such as requiring additional devices other than a mobile device or resulting in large computational complexity because of the use of camera images for gesture recognition.

This paper proposes hand gesture interface using light-dark changes in an illuminance meter built in a mobile device (hereinafter referred to as "built-in illuminance meter") or "HGI/LI" in short, and examines HGI/LI's precision in hand gesture recognition. It also evaluates usability of HGI/LI by conducting an experiment with human subjects.

2. Hand gesture interface using mobile devices

Active researches have been conducted on hand gesture interface using mobile devices in recent years. SideSwipe [9] detects and recognizes hand gestures using GSM signal by a circuit board with four antennas attached to the back of a smartphone. This realizes recognition of hand gestures not only over the device but also those around it. SurfaceLink [10] recognizes gestures using an accelerometer, a vibration motor, a speaker, and a microphone built in smartphones. This realizes sharing and exchanging information among multiple devices on the same surface.

In these studies, however, additional devices other than mobile devices are required in order to recognize hand gestures. Therefore, if additional devices are expensive, the introduction cost of the system is high. Even if additional devices are inexpensive, a method using mobile devices only is preferred in light of advantages of using mobile devices which have become generic products, although it depends on the types of gestures recognized, the precision of recognition, and applications realized by a given method. HGI/LI proposed in this paper uses a single mobile device only in order to recognize hand gestures.

Song et al. [11] expanded the interaction with a mobile device by using an RGB camera built in a smartphone to recognize gestures and combining this with touch operation. Robust gesture recognition is realized by using an algorithm based on the random forest. Recognizing hand gestures using camera images, however, requires large computational complexity in analyzing images. The battery duration of a mobile device needs to be considered. Privacy consideration is also required in using camera images.

HGI/Li uses a built-in illuminance meter to recognize hand gestures with small computational complexity. This can realize lighter applications. Unlike a method using camera images, it does not require privacy consideration.

3. Hand gesture interface using light-dark changes in an illuminance meter built in mobile devices

3.1 Concept

A diverse variety of sensors, including proximity sensor and accelerometer, are installed on smartphones and tablets. An illuminance meter is also built in a mobile device for the purpose of adjusting the brightness of its display. This study proposes HGI/LI, which realizes a hand gesture interface using light-dark change of the built-in illuminance meter. HGI/LI realizes intuitive operation by hand gestures on a mobile device by obtaining illuminance information from an illuminance meter, extracting features from changes in illuminance values, and classifying gestures into five types using a classification model obtained through decision tree learning.

3.2 Type of hand gesture

HGI/LI recognizes five types of hand gestures: hide, roll, up, down, and slash. Fig. 1 - 5 is a picture of these hand gestures.

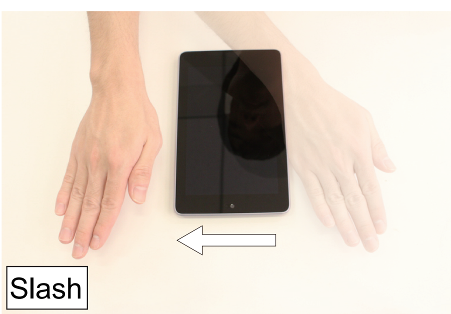

Fig. 1: Slash gesture

Fig. 2: Roll gesture

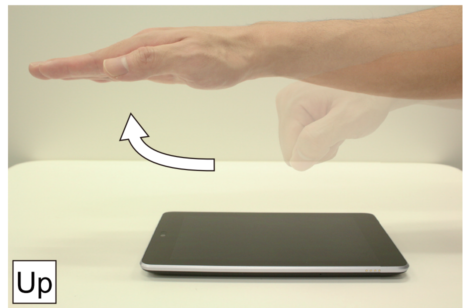

Fig. 3: Up gesture

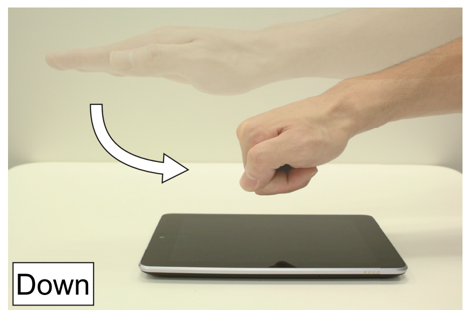

Fig. 4: Down gesture

Fig. 5: Hide gesture

A "roll" gesture is a motion of turning your hand around over the built-in illuminance meter. A "slash" gesture is a motion of moving your hand in horizontally. An "up" gesture is a motion of moving your hand up while pushing it out. A "down" gesture is a motion of pulling your hand toward you while moving it down. A "hide" gesture is a motion of hiding the illuminance meter with your hand. It causes a large change in the illuminance value relative to other gestures, resulting in the illuminance value of almost 0.

3.3 Hand gesture recognition algorithm

The hand gesture recognition algorithm is given below, and the explanation of each step follows.

1) Detect an illuminance change by the built-in illuminance meter and obtain illuminance values.
2) If an illuminance change greater than the threshold occurs, obtain the set of data from that point to the point where illuminance values stabilize at the original value again.
3) Extract features from the data set obtained and classify the corresponding motion as the operation mode toggling gesture or disturbance.
4) If it is classified as the operation mode toggling gesture, enter the operation mode and move to step (5). If it is classified as disturbance, go back to step (1).
5) Just as in step (1), detect an illuminance change by the built-in illuminance meter and obtain illuminance values.
6) Just as in step (2), if an illuminance change greater than the threshold occurs, obtain the set of data from that point to the point where illuminance values stabilize at the original value again.
7) Extract features from the data set obtained and classify the corresponding motion as the operation mode toggling gesture or other hand gestures.
8) If it is classified as the operation mode toggling gesture, exit the operation mode and go back to step (1). If it is classified as other hand gestures, execute the process assigned to each hand gesture.
9) Repeat steps (5) through (8) until the operation mode toggling gesture is recognized.

First of all, HGI/LI detects an illuminance change by the built-in illuminance meter and obtains an illuminance value. In obtaining illuminance values, values obtained by the sensor fluctuate even if the lighting luminance is kept constant. This issue is resolved by setting a threshold based on the result of a preliminary experiment. If a change in illuminance values is below the threshold, the current state is judged to be such that the lighting luminance is kept constant and that no hand gesture is made. If a change in illuminance values is greater than the threshold, a set of data is to be obtained from that point to the point where illuminance values stay constant again. Features are extracted from the data set obtained, and hand gesture recognition is performed by using extracted features and classification models obtained in advance through machine learning. Feature extraction and the hand gesture classification method in HGI/LI are elaborated in the next subsection.

Let us now explain steps (3), (8), and (9). In performing hand gesture recognition using light-dark change, it is necessary to consider disturbance, just as in other methods. An illuminance meter detects a light-dark change due to the tilt of the device and also affected by the shadow of a person or papers. In order to enhance the precision in hand gesture recognition, this interface has the operation mode. The operation mode corresponds to steps (5) through (9). Operations by other hand gestures are enabled after entering the operation mode using the operation mode toggling gesture. In this study, a "hide" gesture is adopted as the operation mode toggling gesture. Since performing an operation by a "hide" gesture causes a large change in illuminance values, resulting in the illuminance value of almost 0, it is considered possible to distinguish a "hide" gesture not only from other hand gestures but also from disturbance.

In HGI/LI, it is also necessary to consider a change in the light environment. A change in the light environment occurs when a user moves to a room whose light environment is different or when the lighting luminance is changed in an environment in which a controllable lighting system is used. A light environment herein refers to an illuminance environment. This issue can be dealt with by resetting the data set subject to detection when a change in illuminance values stabilizes at or below the threshold for a specified time or longer and by going back to step (1) from step (2) or to step (5) from step (6).

3.4 Feature extraction and classification of hand gesture

Performing an operation by hand gesture causes a change in illuminance values given by a built-in illuminance meter to generate a wave of illuminance values. This wave is obtained as a data set, from which features are extracted.

HGI/LI classifies hand gestures by extracting the total of four features: the number of waves of illuminance values and features D, S, and Tt given by equations (1) through (3).

$$D = \frac{A}{I} \tag{1}$$

$$S = \left|\frac{A}{T_s}\right| - \frac{A}{T_e} \tag{2}$$

$$T_t = T_s + T_e \tag{3}$$

A : amplitude, I : illuminance environment [lx]
T_s : time from start point to deepest point of wave [ms]
T_e : time from deepest point to end point of wave [ms]

Decision tree learning is used for classification. A decision tree is composed of root node, split nodes, and leaf nodes. Classification starts at the root node, and each split node classifies an input value into one of children nodes based

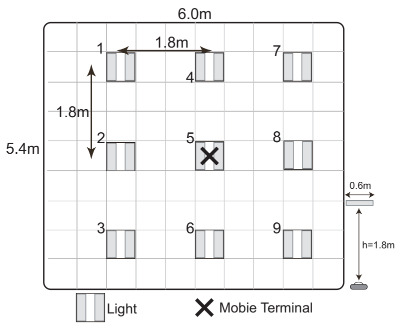

Fig. 6: Experimental environment

Fig. 7: Experimental situation on the recognition accuracy

on the result of learning. Classification is performed by repeating this until reaching leaf nodes.

While decision tree learning is not a method that has high classification precision, it has characteristics that it is highly accessible to human understanding and highly readable. It is one of most widely used learning methods. It also has such characteristics that, in comparison with other methods, it requires less computational complexity in performing classification and enables a faster classification. This study attempts to recognize hand gestures with less computational complexity by classifying them using a shallow decision tree based on four features mentioned above.

4. Hand gesture recognition accuracy verification

4.1 Experimental overview

A verification experiment was conducted on HGI/LI's hand gesture recognition precision. Fig. 6 shows an experiment environment, and Fig. 7, a scene of experiment. The experiment was conducted by placing a desk directly under the lamp in the center of nine LED lamps and a mobile device on the surface of the desk. The installation interval of lighting fixtures was 1.8 m, which is the same as that in a typical office. Subjects were 7 students aging from 23 to 24. The experiment was conducted after subjects were briefed for about 5 minutes about each gesture before the experiment.

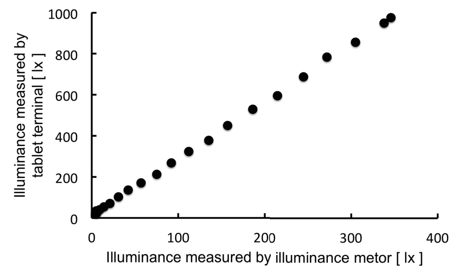

Fig. 8: Comparison of illuminance values mesured by mobile terminal and illuminance metor

In the experiment, subjects performed five types of gestures indicated in Chapter 3. Four patterns of illuminance environment were prepared, and the experiment was conducted by changing the illuminance on the desk surface to 1000 lx, 700 lx, 500 lx, and 300 lx. This was repeated 10 times in total. These 1400 data were collected to evaluate precision verification in order to avoid a bias in data resulting from learning or fatigue, instructions for gestures were randomly given and the illuminance environment was randomly changed in the experiment.

In conducting the recognition precision verification experiment, a performance verification experiment was conducted with a mobile device to be used in the former as a preliminary experiment.

4.2 Performance verification of built-in illuminance meter

As performance verification experiments of built-in illuminance meters, we conducted an experiment comparing values obtained by the built-in illuminance meter and illuminance values measured by an illuminance meter and one evaluating the response performance of the built-in illuminance meter. A mobile device used in this study is Nexus 7 (2012 model) tablet.

In the experiment comparing values obtained by the built-in illuminance meter and measured values of the illuminance meter, the illuminance meter and the mobile tablet were placed on the desk surface and a single lamp directly above them was turned on. A light fixture with dimming control in 256 levels was used in the experiment. Brightness at each step was measured by using the built-in illuminance meter and the illuminometer. As for an illuminometer, ANA-F11 made by Tokyo Koden was used. Fig. 8 shows the comparison of values obtained by the built-in illuminance meter and measured values of the illuminometer. The number of plots shown are reduced in order to make the result easier to see. As a result of the experiment, while values obtained by the built-in illuminance meter were greater than measured values of the illuminometer, the linearity of their relation was confirmed.

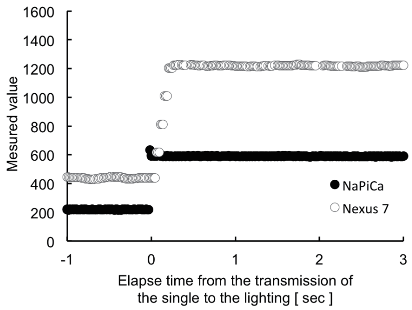

Fig. 9: Reaction performance of a built-in illuminance sensor

The response performance verification experiment of the built-in illuminance meter examined the interval at which the built-in illuminance meter obtains an illuminance value and time required for illuminance values to converge to the correct value. Just as in the experiment described above, this experiment was conducted by placing the mobile device on the desk surface and turning on only a lamp directly above.

In the experiment, time from the transmission of a light control signal to the lamp to the convergence of values obtained by the built-in illuminance meter to a constant value was measured by first turning on the lamp at 30% of the maximum lighting luminance and then raising luminance to 90% of the maximum lighting luminance. In order to measure a precise change in illuminance, illuminance values were measured, at the same time, by using NaPiCa illuminance meter made by Panasonic [12]. This meter can obtain illuminance values at an interval of approximately 1.2 ms.

The result of the experiment is shown in Fig. 9. The horizontal axis indicates the time elapsed from the time a light control signal was transmitted to the lamp, and the vertical axis indicates values obtained by the built-in illuminance meter. As a result of the experiment, it was found that the built-in illuminance meter obtained illuminance information at an interval of approximately 50 ms and that approximately 200 ms of time was required from the transmission of a light control signal to the lamp to the convergence of illuminance values to a constant value.

4.3 Experimental Results

Evaluation was done by four patterns: leave-one-out cross validation (LOOCV) using data for al subjects, LOOCV using data for each subject, leave-one-subject-out cross validation (LOSOCV), and leave-one illuminance-out cross validation (LOIOCV).

LOOCV using data for each subject, for which both test and training data are composed of data for one subject only to evaluate classification precision for each subject. LOSOCV groups data by subjects and uses data for one subject as test data and data for other subjects as training data. Generic applicability to unknown users is evaluated by using LOSOCV in evaluation. LOIOCV groups data by illuminance environments and uses data for one illuminance environment as test data and data for other illuminance environments as training data. Generic applicability to unknown environments is evaluated by using LOIOCV in evaluation.

Table 1 gives the result of LOOCV using data for all subjects, and Table 2, other results. Table 1 shows the recognition rate of all hand gestures by confusion matrix. "Per User" in Table 2 refers to the result of LOOCV using data for each subject.

Table 1: Confusion matrix for 5 recognized gestures(LOOCV)

	Hide	Roll	Up	Down	Slash
Hide[%]	96.8	0.0	2.1	1.1	0.0
Roll[%]	0.0	95.4	2.9	1.4	0.4
Up[%]	0.7	0.0	94.3	0.4	4.6
Down[%]	1.4	0.0	1.4	93.6	3.6
Slash[%]	0.0	0.0	3.2	0.4	96.4

Table 2: Results of Per User, LOSOCV, and LOIOCV

	Hide	Roll	Up	Down	Slash	Average
Per User[%]	97.5	98.6	91.8	98.2	96.4	96.5
LOSOCV[%]	95.4	96.1	90.0	96.8	91.8	94.0
LOIOCV[%]	96.1	98.2	94.3	97.9	95.4	96.4

Based on Table 1, it is found that HGI/LI shows high precision of 95.3% on average and recognizes hand gesture of each type at high precision of 93.6% or greater. In addition, based on Table 2, it is found that gestures were classified at precision of 91.8% or greater, 90% or greater, and 94.3% or greater according to evaluation by Per User, LOSOCV, and LOIOCV, respectively.

4.4 Discussion

Fig. 10 shows the decision tree created by using 1,400 data collected in this experiment. Let us dwell on what causes hand gesture recognition precision to decline in HGI/LI.

With the tablet used in this study, it takes approximately 200 ms to converge to a correct value. In addition, Fig. 10 shows that gestures are classified into "hide" gestures and other gestures by using the feature D. Consequently, when a gesture is executed so fast that it is finished before the illuminance value fully declines, a "hide" gesture may be wrongly recognized as another gesture.

Table 1 tells that the probability of wrongly recognizing "up" and "down" gestures as "slash" gestures is higher than those of other patterns of misrecognition. Based on Fig. 10, it is confirmed that HGI/LI classifies gestures into "up," "down," and "slash" gestures mostly by using the feature Tt. Therefore, this is considered as misrecognition caused by

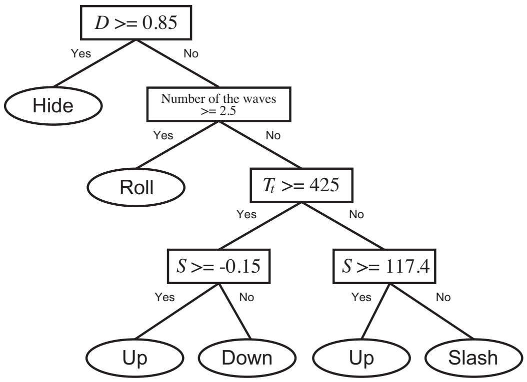

Fig. 10: Dicision tree with 1400 data collected in the experiment on recognition accuracy(D, S, and T_t is explained in Section 3.4 as Expression 1 - 3)

the fast execution speed of some "up" or "down" gestures. A "slash" gesture executed too carefully results in its slower speed, which is considered to cause its misrecognition for "up" or "down" gesture.

While LOSOCV resulted in a lower recognition rate than the rate evaluated by LOOCV using data for each subject, it still shows high precision at 94.0% on average. Based on this result, HGI/LI is considered to have high generic applicability to unknown users. By LOSOCV, however, the recognition rate was 80.5% on average over all gestures. In particular its recognition rate of "slash" gestures was low at only 60.0%.

On the other hand, LOOCV resulted in the recognition rate of 95.4% on average for all gestures and the recognition rate of 96.4% for "slash" gestures. Since high recognition rates are given by LOOCV, including data of the user who actually uses the mobile device in learning data is considered to ensure a high recognition rate.

As LOIOCV resulted in the high recognition rate of 96.4% on average, operations by hand gestures through HGI/LI are considered to be possible if the illuminance value of the environment in which HGI/LI is used ranges between 300 lx and 1000 lx.

5. Usability evaluation

5.1 Experiment overview

This section describes a subjective experiment conducted with 8 students aging from 22 to 24 to evaluate the usability of each hand gesture in HGI/LI. A questionnaire survey and an interview were conducted with subjects after the experiment to provide materials for reflection on HGI/LI.

In this experiment, a demonstration of about five minutes was first given to subjects. We then had subjects actually use HGI/LI and conducted a subjective questionnaire survey about each gesture. The questionnaire was composed of 5-point Lickert Scale questions and asked respondents to evaluate the ease of each gesture. The experiment environment is the same as that of the precision verification experiment described in the previous subsection. The illuminance environment is set to 700 lx. The experimental environment is shown in Fig. 11.

Fig. 11: Experimental situation on the usability evaluation

5.2 Application

In order to evaluate usability, we implemented HGI/LI and created an application displaying the result of classification of a gesture on a PC display. In this application, a "hide" gesture was adopted as an operation mode toggling gesture. We therefore conducted a preliminary experiment for classifying gestures into "hide" gestures and other disturbance before creating the application. The preliminary experiment was conducted in the experiment environment shown in Fig. 6.

In the preliminary experiment, disturbance data affecting gesture recognition were collected first with cooperation by three subjects. In this experiment, illuminance changes occurring upon the following motions were collected as disturbance data.

- Look into the mobile device
- Stand up and sit down
- Leave the seat
- Sit the seat
- Move along the table
- Move the paper above the mobile device
- Tilt the mobile device

We had three subjects repeat the above motions five times, changing the speed of each motion each time. Accordingly, 105 data were collected. Then, learning was done using "hide" gesture data collected in the illuminance environment of 700 lx, as described in the previous section, and these disturbance data to create a decision tree.

This application classifies gestures into operation mode toggling gestures and other disturbance data by using this decision tree. The application created also classifies gestures by using the decision tree created on the basis of 1,400 data shown in the previous experiment.

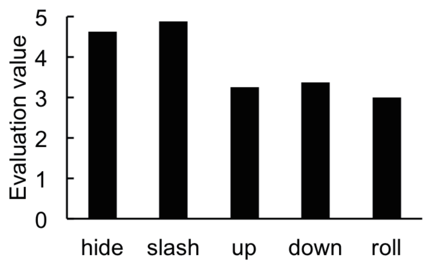

Fig. 12: Evaluation for easiness of each gestures

5.3 Results and Consideration

Fig. 12 gives the result of the questionnaire survey about the ease of each gesture. Fig. 12 shows that, while "hide" and "slash" gestures obtained high scores, the scores for other three kinds of gestures were below 4. On "up" and "down" gestures, there were remarks such as "It is a bit hard to move my elbow, wrist, and fingers at the same time," and "I find back-and-forth motions hard such as pulling my hard toward me, pushing my hand forward." On a "roll" gesture, there were remarks such as "Interaction time is long," and "A motion that includes not only horizontal but also vertical movements is bothersome." A "roll" gesture is a motion such that you move your hand in a circle across the built-in illuminance meter back-and-forth 1.5 times. This gesture is distinguished from other gestures by using the number of waves of illuminance values. Therefore, it is misrecognized for other gestures unless the user’s hand moves across the built-in illuminance meter three times. Multiple subjects felt this gesture difficult. In the future, it is considered necessary to evaluate usability again after creating a specific application using HGI/LI and conducting an experiment that compares it with other interfaces for the same operations.

6. Conclusion

This study examined HGI/LI, hand gesture interface using light-dark changes in the illuminance meter built in a mobile device. While studies on gesture interface often uses an in-frared camera or depth sensor to recognize gestures, doing so requires a dedicated device. As HGI/LI uses the illuminance meter built in a mobile device, which has rapidly spread in recent years, its introduction cost can be reduced. In addition, HGI/LI minimizes computational complexity and realizes a lighter application as it uses light-dark changes in an illuminance meter and perform classification only by using a shallow decision tree.

As a result of the experiment for verifying the hand gesture recognition precision, HGI/LI showed recognition rates of 93.6% or above for all gestures. It was found that hand gestures can be classified into five types using a shallow decision tree based on four types of features extracted from illuminance information. The generic applicability of HGI/LI to unknown users and illuminance environments was also confirmed.

References

[1] Cohn G. and Morris D. and Pate S. and Tan D., "Humantenna: Using the Body as an Antenna for Real-time Whole-body Interaction,", "In Proc CHI 2012, pp.1901-1919
[2] Gupta S. and Morris D. and Patel S. and Tan D., "Soundwave: Using the Dopper Effect of Sense Gestures, "In Proc CHI 2012, vol.1, no.1322, pp.1911-1914
[3] Harrison C. and Tan D. and Morris D., "Skinput: Appropriating the Body as an Input Surface, "In Proc CHI 2010, pp.456-462
[4] Shotton J. and Fitzgibbon A. and Cook M. et al, "Real-time human pose recogniition in parts from single depth images, "In Proc CVPR 2011, pp.189-192
[5] Leap Motion, https://www.leapmotion.com/
[6] Lee J. and Olwal A. and Ishii H. and Boulanger C., "SpaceTop: Integrating 2D and Spatial 3D Interactions in a See-through Desktop Environment, "In Proc CHI 2013, pp.189-192
[7] Taylor S. and Keskin C. and Hilliges O. and Izadi S. and Helmes J., "Type-hover-swipe in 96 Bytes: A Motion Sensing Mechanical Keyboard, "In Proc CHI 2014, pp.1695-1704
[8] Juan W. and Helman S. and Yael E. and Michael G. and Craig F. and Mark S. and Jon H., "A Real-Time Hand Gesture Interfece for Medical Visualization Applications, "Applications of Soft Computing, pp.153-162
[9] Zhao C. and Chen K. and Aumi M. T. I. and Patel S. and Reynolds M. S., "SideSwipe : Detecting in-air Gestures Around Mobile Devices Using Actual GSM Signals, "In Proc UIST 2014, pp.527-534
[10] Goel M. and Lee B. and Aumi I. T. Md. and Patel S. and Borriello G. and Hibino S. and Begole J., "SurfaceLink: Using Inertial and Acoustic Sensing to Enable Multi-device Interaction on a Surface, "In Proc CHI 2014, pp.1387-1396
[11] Song J. and Sörös Gábor and Pece F. and Fanello S. R. and Izadi S. and Keskin C. and Hilliges O., "In-air Gestures Around Unmodified Mobile Devices, "In Proc UIST 2014, pp.319-329
[12] Panasonic:illuminance meter NaPiCa, http://www3.panasonic.biz/ac/download/control/sensor/illuminance/catalog/bltn_jpn_ams.pdf.