

Article

Multi-touch Tabletop System Using Infrared Image Recognition for User Position Identification

Shota Suto ¹, Toshiya Watanabe ², Susumu Shibusawa ^{3,*} and Masaru Kamada ²

¹ Graduate School of Science and Engineering, Ibaraki University, Hitachi, Ibaraki 316-8511, Japan; East Japan Institute of Technology Co., Ltd., Hitachi, Ibaraki 319-1221, Japan; macnamas@yahoo.co.jp

² College of Engineering, Ibaraki University, Hitachi, Ibaraki 316-8511, Japan; toshiya.watanabe.inu@vc.ibaraki.ac.jp; masaru.kamada.snoopy@vc.ibaraki.ac.jp

³ Ibaraki University, Hitachi, Ibaraki 316-8511, Japan; susumu.shibusawa.ptr@vc.ibaraki.ac.jp

* Correspondence: susumu.shibusawa.ptr@vc.ibaraki.ac.jp; Tel.: +81-294-38-5153

Abstract: A tabletop system can facilitate multi-user collaboration in a variety of settings including small meetings, group work, and education and training exercises. The ability of identifying the users touching the table and their positions can promote collaborative work among participants, so methods have been studied that involve the attaching of sensors to the table or chairs or to the users themselves. An effective method of recognizing user actions without placing a burden on the user would be some type of visual process, so the development of a method that processes multi-touch gestures by visual means is desired. This paper describes the development of a multi-touch tabletop system using infrared image recognition for user position identification and presents the results of touch-gesture recognition experiments and a system usability evaluation. Using an FTIR touch panel and infrared light, this system picks up the shadow area of the user's hand by infrared camera in relation to user touch operations and estimates user position by image recognition. The multi-touch gestures prepared for this system include an operation to change the direction of an object to face the user and a copy operation in which two users generate duplicates of an object. The system usability evaluation revealed that prior learning was easy and that system operations could be easily performed.

Keywords: tabletop system; user position identification; infrared image recognition; multi-touch gesture; FTIR panel; system usability

1. Introduction

Hand gestures are a natural form of human communication and are seen as a promising means of human-computer interaction [1], [2]. A tabletop system that allows for direct-touch input enables input by both hands in a natural and smooth manner surpassing conventional mouse and keyboard devices. It is expected to decrease the user's cognitive load in interacting with content [3]. In small meetings and group work, a tabletop system is expected to provide an environment conducive to collaborative work. That is, in face-to-face interaction in which multiple users gather around a table, we can expect such a system to improve the contributions of each user and generate a sense of teamwork [4].

There have been many research studies to date in relation to tabletop systems including touch screen technology [5], a method for achieving an interactive table [6], research on multi-touch gestures [7], [8], and application to collaborative work support [4], [9]. Among the methods used to achieve an interactive table, the frustrated total internal reflection (FTIR) method can obtain infrared images through a simple and inexpensive mechanism that irradiates the interior of an acrylic panel with infrared light and detects leaked light from the section of the panel touched by the user using an infrared camera [6].

In the case of a multi-touch table used for simultaneous interaction by multiple users, being able to identify the users who are touching the table and their positions can facilitate group work, education and training [4], medical training and treatment planning [10], security control for machines, games, etc. that take participants into account. To therefore obtain information on users, parts of their bodies, user positions, etc., methods for attaching sensors to the table or chairs or to the users themselves [11], [12], [13] and a method based on visual processing using a camera [14] have been studied.

A method for identifying users by attaching sensors to the table or chairs or to people requires time to install those sensors and a step for learning how to use them. It may also place constraints on user position and posture. In contrast, a visual method that detects parts of a user's body by image recognition technology to recognize user actions is one way of constructing a system that does not require a sensor to be attached to one's body thereby decreasing the burden on the user. A tabletop system capable of user identification requires the recognition of both users and multi-touch gestures, so further progress in the research of multi-touch gestures by visual methods has been desired.

We have previously created a tabletop system using infrared image recognition for user position identification [15], [16]. In this paper, we describe a multi-touch tabletop system using infrared image recognition for user position identification that expands upon the previous system configuration and software. We also present the results of touch-gesture recognition accuracy experiments and a system usability evaluation.

This tabletop system uses an infrared light placed above the FTIR multi-touch panel and an infrared camera to obtain information on touch points and the shadow area of the hand when a user performs a touch operation. The system uses this information on touch points and the shadow area of the hand to establish an association between the hand and table touch points and estimate the position of the user touching the table. The multi-touch gestures prepared for this system include an operation to change the direction of an object to face the user and a copy operation in which two users generate copies of an object. The system usability evaluation was conducted on the basis of a questionnaire based on the System Usability Scale (SUS) evaluation method [17] to which we have added a section for free comments.

The rest of this paper is organized as follows. Section 2 summarizes related research on tabletop systems, Section 3 describes the proposed system and the user-position estimation technique, Section 4 describes an actual implementation of the system, Section 5 presents and discusses the results of touch-gesture recognition accuracy experiments and a usability evaluation using the implemented system, and Section 6 summarizes this study and touches upon future issues.

2. Related Research

This section describes previous research related to user collaboration support by tabletop and its application, tabletop sensing methods, multi-touch gestures, and user position identification.

2.1. User Collaboration Support by Tabletop and Its Application

Interaction among multiple users using a tabletop system can be broadly divided into face-to-face interaction around a single table and distributed interaction around tables installed in different spaces. A tabletop system based on face-to-face interaction can help each user make a deeper contribution to the topic of discussion while promoting team bonding. Systems of this type are expected to be especially effective in educational activities for young people and the planning of medical treatment through interaction between medical personnel.

Morris et al. [4] used the software of the DiamondTouch table capable of multi-user identification to implement 11 gesture applications and study the use of cooperative gestures in multi-user interaction. Evans et al. [9], meanwhile, examined the relationship between touch interactions and the collaborative process in field studies of adolescent students and showed that touch patterns reflect the quality of collaboration. In addition, Ohashi et al. [18] constructed a computerized KJ method support system enabling finger pointing, measured working time and

number of comments made, and showed that this system with finger pointing could reduce working time.

In the fields of clinical medical treatment and surgery, there are high hopes for practical technologies based on mixed reality that overlay information from diverse types of sensors with real-time images. The advancement of interaction technologies in these fields has therefore become a major issue [19]. Lundström et al. [10][10] developed a table system for medical visualization for application to orthopedic surgery planning and discussed issues in system design. They found that essential design objectives in the configuration of such a system that includes interaction are high similarity with actual physical conditions and a very low learning threshold.

Genest et al. [20] developed a toolkit called KinectArms to capture and display arm embodiments with the aim of facilitating gesture-driven communication in remotely distributed tables. KinectArms provides a visual representation of arms by using a depth camera to determine gesture height and improves the expressive power and usability of distributed tabletop groupware.

2.2. Tabletop Sensing Methods

A number of touchscreen technologies have been developed to enable a person to manipulate a display screen through touch. These include projected capacitive, analog resistive, infrared, camera-based optical, planer scatter detection, vision-based, and combinations of technologies. Walker [5] presented a broad overview of 13 types of touchscreen technologies describing for each a brief history, basic operating principle, typical applications, main advantages and disadvantages, current issues and trends, and future outlook. He described the projected capacitive method in more detail than the other methods due to its current dominance.

Han [6] described a detailed implementation of an FTIR-based multi-touch, interactive table and outlined the future direction of multi-touch sensing technology. Zhang et al. [21] introduced sensing technology that enables touch input on the surface of objects having irregular and complex forms using electric field tomography and demonstrated the feasibility of this technology using example applications.

2.3. Multi-Touch Gestures

Typical gestures on a tabletop include move, zoom in/out, rotate, drag, tap, flick, and hold. The tabletop user uses the gestures created by the system designer. These gestures, though thought to be appropriate in initial research, do not necessarily reflect user actions. Designing natural and intuitive gestures for a new multi-touch interface therefore requires a survey on how users would approach a multi-touch interface and the types of gestures they would use.

North et al. [22] asked users to execute an object-sorting task on a physical table, multi-touch surface, and desktop computer with a mouse, measured and compared task execution times, and collected the set of user gestures on the multi-touch surface. Wobbrock et al. [7] collected 1080 gestures for the case of 20 non-expert users using a tabletop with one and two hands, paired those gestures with 27 commands, and examined the way in which users used multi-touch gestures. Hinrichs et al. [8] used a large multi-touch tabletop exhibited at a municipal aquarium to conduct a field survey on the use of multi-touch gestures by visitors. They found that the use of multi-touch gestures was influenced by user preference, usage conditions, and social conditions and that previous gestures influenced subsequent gestures to form gesture sequences.

2.4. User Position Identification

The ability of identifying the users touching a multi-touch tabletop and their positions opens the door to diverse methods of use in collaborative group work and other scenarios. Dietz et al. [11] described the design method, construction, and usage results of DiamondTouch, a technology for identifying the positions of particular users touching a multi-user touch table from electric fields generated by capacitive coupling between the users and their chairs. Marquardt et al. [12] developed the TouchID toolkit for multi-touch tabletop interaction introducing fiduciary-tagged gloves and

described its suite of techniques. This toolkit can gather information on the person touching the table, the hand being used, which hand part, and hand posture and gesture.

Annett et al. [23][23] created a tabletop system equipped with 138 proximity sensors around a Microsoft Surface to detect a user's position, distinguish between left and right arms, and establish a correspondence between touch points and users and hands. Lissermann et al. [13] created an environment supporting group work, individual work, and in-between transitions using a multi-view tabletop consisting of a multi-touch frame, 3D display, two Kinect cameras for user and hand recognition, and 3D shutter glasses worn by users. They described its implementation techniques and presented application examples. Zhang et al. [14] determined the contours of users' hands using an infrared lamp above an FTIR table, predicted user position by machine learning based on the finger orientation distributions of users touching the tabletop surface, and measured the accuracy achieved.

Finally, Suto et al. [15], [16] created a multi-touch tabletop system that identifies user position by image recognition using an FTIR touch panel and external infrared light. They investigated the accuracy of recognizing multi-touch gestures with this system.

3. System Configuration

3.1. System Overview

The basic configuration of a multi-touch tabletop system consisting of an FTIR table and infrared light is shown in Figure 1. This system installs an infrared camera underneath the table to capture the acrylic panel on top of the table and a projector connected to a personal computer (PC) to project images onto the panel. It also installs an infrared floodlight consisting of an infrared LED on both ends of the acrylic panel for irradiating the panel with infrared light. In addition, the system pastes tracing paper that plays the role of a screen onto the acrylic panel that becomes the image-projection surface and presents information to users by projecting images from the projector.

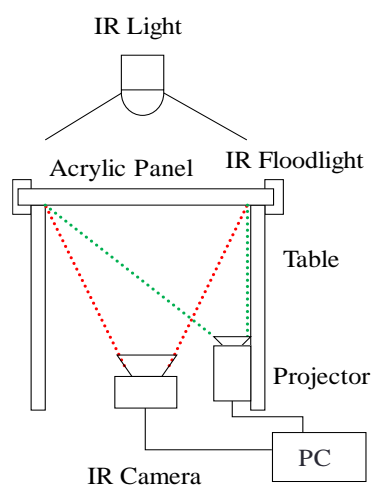


Figure 1. System configuration.

Furthermore, to obtain information on a user's hand area, the system installs an infrared light on the ceiling above the table. Since a user's hand on the tabletop will obstruct infrared beams from this light, an infrared shadow corresponding to the area of the hand will form enabling the shadow of this hand to be picked up by the infrared camera.

3.2. Overview of User-Position Estimation Technique

In a tabletop system, a user extends a hand from the edge of the table to manipulate an object displayed on the tabletop. At this time, the ability of determining from which direction the hand touching the object is being extended would make it possible to estimate the position of the user manipulating those touch points.

Touch points on the FTIR table appear as white light and the shadow of the user's hand appears on the table owing to the overhead infrared light. The infrared camera picks up both of these images. Now, by performing background differencing to these captured images, the tabletop image can be classified into three types of areas: the touch areas having higher brightness values than the background, the hand area having lower brightness values than the background, and the background itself. Here, brightness value b of the captured tabletop image can be expressed as follows with respect to threshold values σ_1, σ_2 ($\sigma_1 < \sigma_2$):

$$\left. \begin{array}{l} b < \sigma_1 \quad : (\text{Hand Shadow}) \\ \sigma_1 \leq b < \sigma_2 \quad : (\text{Background}) \\ b \geq \sigma_2 \quad : (\text{Touch Area}) \end{array} \right\} \quad (1)$$

The union of the touch areas and the hand-shadow area constitutes an area having a change in brightness values with respect to the background. It can be defined as the hand area as follows:

$$(b < \sigma_1) \vee (b \geq \sigma_2) : (\text{Hand Area}) \quad (2)$$

The procedure for estimating the position of the user manipulating the touch points is shown in Figure 2. In this process, the system separately extracts the touch points and hand area, superposes and associates these areas, and determines the position of the user manipulating the touch points. A detailed description of this process is described in the following two subsections.

An example of extracting and superposing touch areas and the hand area is shown in Figure 3. The direction of the hand area is determined by the amount of shadow occupying edges of the tabletop. Given that the touch areas constitute a subset of the hand area, the position of the user manipulating the touch points can be estimated. In this example, the system would estimate the user associated with the touch points to be positioned downward relative to the figure.

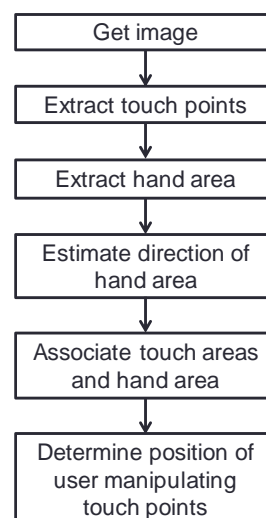


Figure 2. Procedure of estimating position of user manipulating touch points.

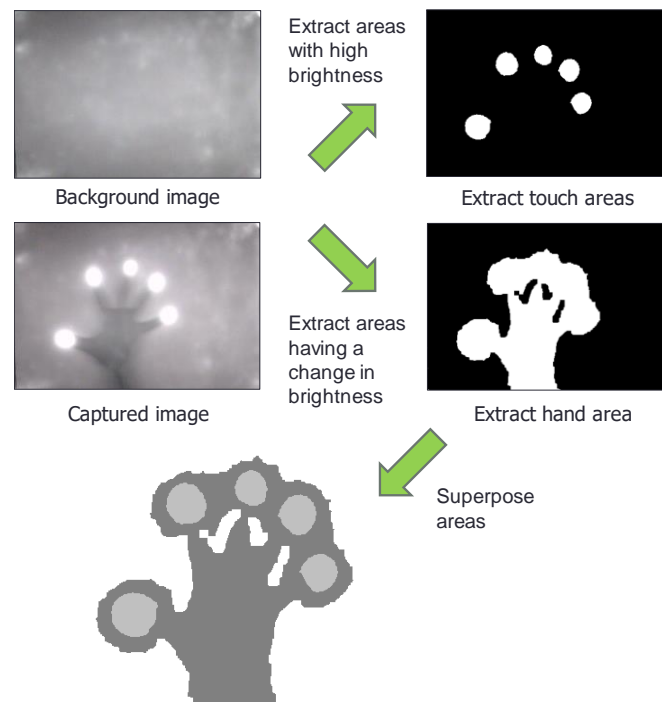


Figure 3. Extraction and superposition of key areas.

3.3. User-Position Estimation Model

3.3.1. Inclusive Relation between Touch Points and Hand Area

In this FTIR system, the touch points are captured as white light and the hand shadow as an infrared shadow generated by the overhead infrared light. The touch areas and hand area can be extracted by performing background differencing on this image. An image of a touch area and that of a hand area including that touch area extracted by background differencing are shown in Figure 4 (a) and (b), respectively. In addition, the positional relationship among the touch area (*TA*), touch point (*TP*), and hand shadow (*HS*) is shown in Figure 4 (c). Here, the outer circle and inner circle at the fingertip corresponds to *TA* and *TP*, respectively. The *TP* is determined by calculating the center of gravity of *TA*.

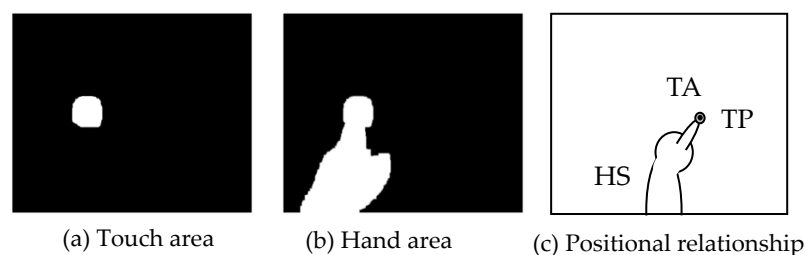


Figure 4. Images of touch area and hand area and their positional relationship.

Given that *TP* is the center of gravity of *TA*, the following relation describing *TP* as an element of *TA* holds:

$$TP \in TA \quad (3)$$

Furthermore, as defined in Eq. (2), hand-area *Hand* is the union of *TA* and *HS*.

$$Hand \equiv TA \cup HS \quad (4)$$

From Eq. (4), the following inclusive relations with respect to *Hand* and *TA/HS* hold:

$$TA \subset Hand \quad (5)$$

$$HS \subset Hand \quad (6)$$

Now, from Eqs. (3) and (5), the following relation holds describing TP as an element of $Hand$:

$$TP \in Hand \quad (7)$$

Although the touch point and hand area are determined by different processing, Eq. (7) shows how a certain touch point relates to a certain hand area.

3.3.2. Model for Estimating User Position

The proposed technique first determines the correspondence between TP and $Hand$. It then determines from which edge in the image the area indicated by $Hand$ is extending and defines that direction as the touch-point direction. The position of the user manipulating that touch point can be estimated in this way.

The model for estimating the position of the user associated with a certain touch point is shown in Figure 5. In the figure, the width and height of the image are denoted as $w+1$ and $h+1$, respectively, and the directions corresponding to the four edges of the tabletop are denoted as d ($d=1, 2, 3, 4$). At this time, the coordinate set $Edge_d$ corresponding to direction d can be expressed as follows:

$$\left. \begin{aligned} Edge_1 &= \{(i,0) | 0 \leq i \leq w\} \\ Edge_2 &= \{(w,j) | 0 \leq j \leq h\} \\ Edge_3 &= \{(i,h) | 0 \leq i \leq w\} \\ Edge_4 &= \{(0,j) | 0 \leq j \leq h\} \end{aligned} \right\} \quad (8)$$

Given the detection of hand-area $Hand$, the intersection of the edge coordinate set and hand area exists, so the following condition holds:

$$Hand \cap \left\{ \bigcup_{d=1}^4 Edge_d \right\} \neq \phi \quad (9)$$

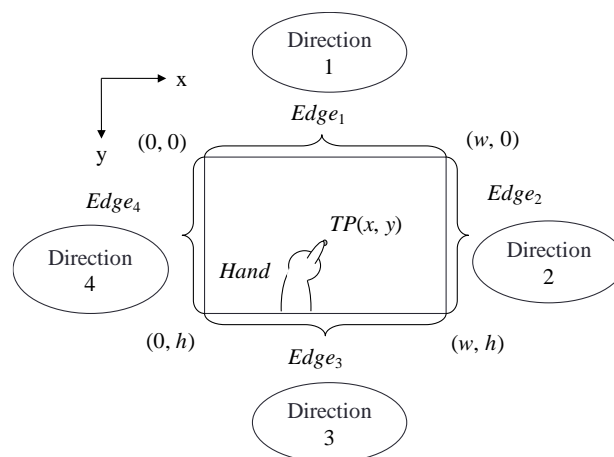


Figure 5. User position estimation model.

3.4. User-Position Estimation Technique

Multiple hand areas usually exist on a multi-user tabletop. To recognize individual hand areas, this technique performs labeling with respect to areas having connected pixels and assigns the label $L[Hand]$ to each hand area $Hand$. It then assigns label $L[TP(x, y)]$ to each touch-point TP in the areas determined by touch-point extraction. According to Eq. (7), $Hand$ includes $TP(x, y)$, which means that $L[TP(x, y)]$ has the same value as label $L[Hand]$ of the hand area manipulating TP .

Referring to Figure 4 (b), *Hand* consists of a contiguous area that may be connected to more than one edge of the image. Accordingly, by investigating which edges *Hand* is actually connected to, the direction from which *Hand* is being extended can be determined. Specifically, by comparing the number of pixels of each edge to which the shadow area of that hand intersects with, the edge with the most intersecting pixels is taken to be the direction from which the hand is being extended. The following steps can be used to estimate the direction of user touch points:

1. Scan label $L[Edge_a]$ of each edge and calculate number of pixels $Pixel_a$ having the same label as hand-label $L[Hand]$.
2. Derive the value of d satisfying $MAX[Pixel_a]$, establish that *Hand* is extending from direction d , and infer that the direction of the position of the user manipulating *TP* is d .

An example of an image of a hand area is shown in Figure 6. Specifically, Figure 6 (a) and (b) show the captured image and the image of the extracted hand area, respectively. In this example, number of pixels $Pixel_a$ of *Hand* on $Edge_a$ takes on a maximum value, which means that *TP* will be taken to be the touch points manipulated by the user positioned in direction 4. We note here that the user in this example is peering down at the tabletop when making gesture operations. This posture results in the casting of a shadow of the user's upper body on the table with the result that the hand area crosses multiple edges.

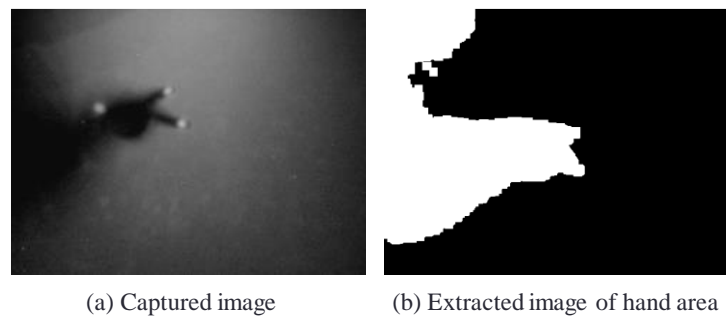


Figure 6. Extraction of hand-area image.

3.5. Object Touch Gestures

This system has a function for manipulating displayed objects through the use of multi-touch gestures. The procedure of object manipulation is shown in Figure 7. In this procedure, the system judges that fingers are touching an object displayed on the tabletop and determines the type of object operation based on the type of finger action. It then executes that object operation and redisplay the object. The touch gestures provided by this system are listed in Table 1 and described below.

- **Move object:**
With one finger touching the object, this gesture moves the object by moving the fingertip. The system detects finger movement and moves the object by only the amount of finger movement in the direction of that movement.
- **Zoom object in/out:**
With two fingers touching the object, this gesture zooms the object in or out by expanding or contracting the space between the fingertips. The system detects the movement of these two fingers and expands the object if that space lengthens and contracts the object if that space shortens.
- **Rotate object:**
With two fingers touching the object, this gesture rotates the object by performing a finger-twisting type of action. The system calculates the angle of rotation from the inclination of the two fingers and rotates the object accordingly.
- **Change direction of object:**
With three fingers touching the object, this gesture changes the direction of the object to face the user. An example of changing the direction of an object by this gesture is shown in Figure 8.

- Copy object:

On judging that two different users are each generating a touch point with respect to a single object, the system duplicates that object. Specifically, in the event that user B performs a single touch on an object while User A is performing a single touch on that object, the object will be copied and placed at the position of User B's touch-point. An example of the copy gesture is shown in Figure 9.

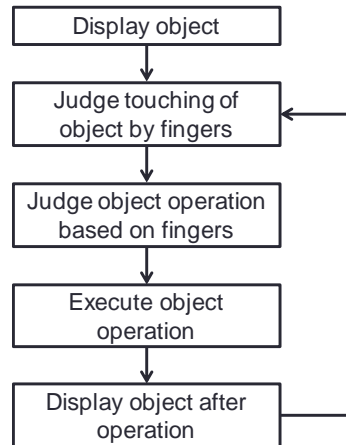


Figure 7. Procedure of object manipulation.

Table 1. Touch gestures.

Operation	No. of Users	No. of Touches	Description
Move	1	1	Move object
Zoom in/out		2	Change object size
Rotate		2	Rotate object
Change direction		3	Change object's direction to face user
Copy	2	2	Copy object

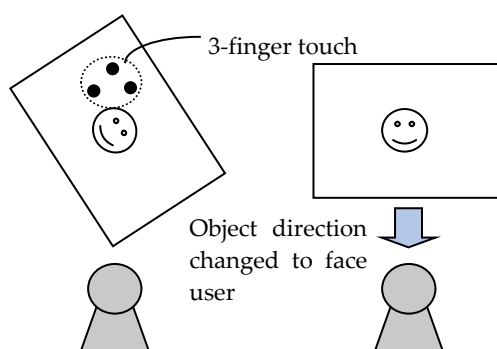


Figure 8. Change-direction gesture.

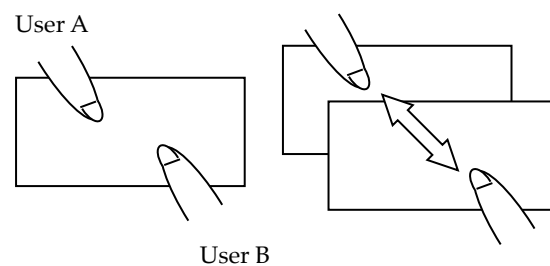


Figure 9. Copy gesture.

4. System Implementation

4.1. Tabletop

We created this system using Visual C and OpenCV running on Microsoft Windows. The FTIR-tabletop infrared floodlight consists of infrared LEDs and a control circuit. The frame rate of the infrared camera is 30 fps maximum. The tabletop itself is 70 cm high with a panel size of 100 cm × 90 cm and a display range of 60 cm × 50 cm. To fix a distance between the projector and the projection

surface, we inserted a mirror between the table and projector. A typical scene of two users manipulating displayed objects using this tabletop system is shown in Figure 10.



Figure 10. Tabletop system.

The procedure of this tabletop system is shown in Figure 11. In this system, the camera pickup area is set somewhat larger than the projector projection area and the rectangular projector projection area is cut out from the camera image to perform image alignment once. As explained in Section 3, obtaining touch points is accomplished by converting the background image and captured image to gray scale, performing difference calculations and threshold processing, removing noise, and extracting and labeling touch areas. Each touch area possesses certain types of information such as center-of-gravity coordinates and number of pixels.

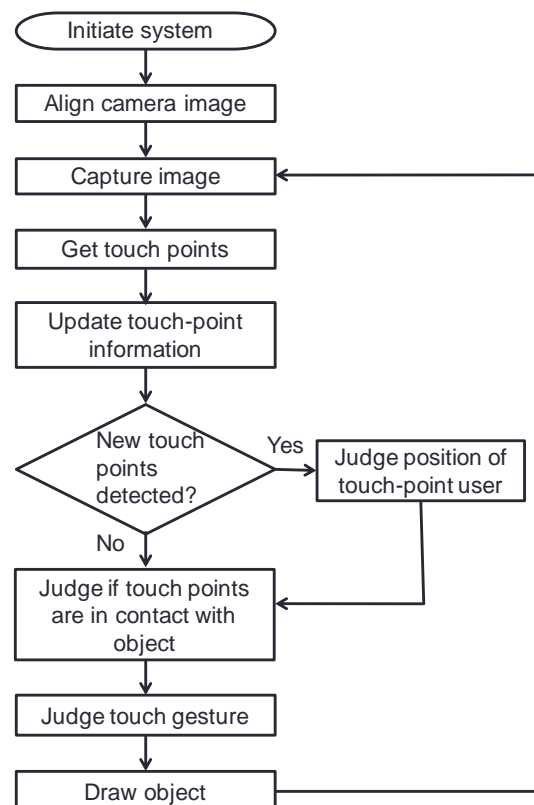


Figure 11. System procedure.

Since touch points appear, move, and disappear by user touch operations, the system compares touch areas between the new and previous frame to update touch-point information. If the centers of gravity of touch areas having the same label should change between the previous frame and new

frame, those touch points are judged to have moved. The process of determining the position of the user manipulating certain touch points follows the design presented in Section 3.

4.2. Photo-Object Manipulation Application

This system is equipped with an application for manipulating photo objects according to the touch gestures performed by multiple users. This photo-object manipulation application reads in image data as photo objects and displays them on the tabletop. The user manipulates objects by touch gestures.

5. Evaluation Experiments

We performed experiments with subjects to measure the recognition accuracy of the change-direction gesture and copy gesture and evaluate the usability of this system.

5.1. Experimental Setup

Taking into account the effects of sunlight on infrared light, we performed the experiments at night. We placed the prototype tabletop described in Section 4.1 in the center of the room and installed two infrared lights on the ceiling above the table spacing them 70 cm apart. Each light was 90-cm long incorporating six equally spaced infrared LEDs. We placed these infrared lights on either side of a fluorescent lamp on the ceiling. The distance from the ceiling to the tabletop panel was 185 cm. The infrared light is shown in Figure 12 and the experimental setup is shown in Figure 13.



Figure 12. Infrared light.

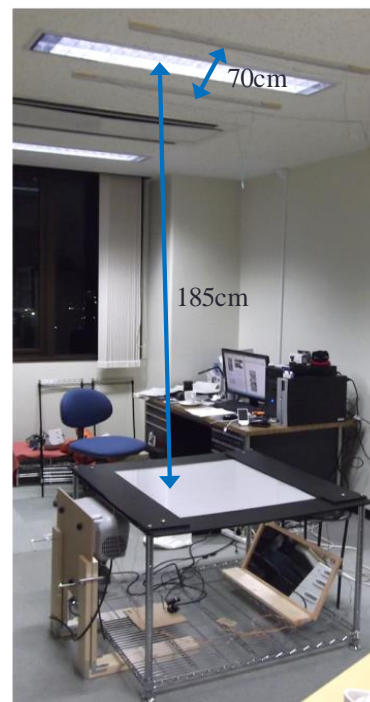


Figure 13. Experimental setup.

5.2. Recognition Accuracy Experiment for Change-Direction Gesture

We performed a subject-based experiment for the 3-finger change-direction gesture and determined the identification rate of user position. Following an explanation of gesture operation, we asked each of four male subjects in their 20s to perform the gesture operation ten times on the tabletop system in each of four different directions. In the experiment, we compared the system-estimated and actual user positions and computed the user-position identification rate. Denoting the number of times the gesture was performed as N_{act} and the number of times the system estimation agreed with the actual user position as $N_{correct}$, the user-position identification rate was computed by Eq. (10).

$$\text{Identification rate} = \frac{N_{\text{correct}}}{N_{\text{act}}} \times 100 [\%] \quad (10)$$

For the sake of clarity, directions $d=1, 3$ and directions $d=4, 2$ in Figure 5 are called up/down directions and left/right directions, respectively. The average identification rate of the change-direction gesture from each of the four directions as performed by the four subjects is shown in Figure 14. The overall average identification rate of the change-direction gesture was approximately 96%. A broad classification of these results reveals that hand direction could be accurately identified in the up/down directions but that there were cases in which it could not be correctly identified in the left/right directions.

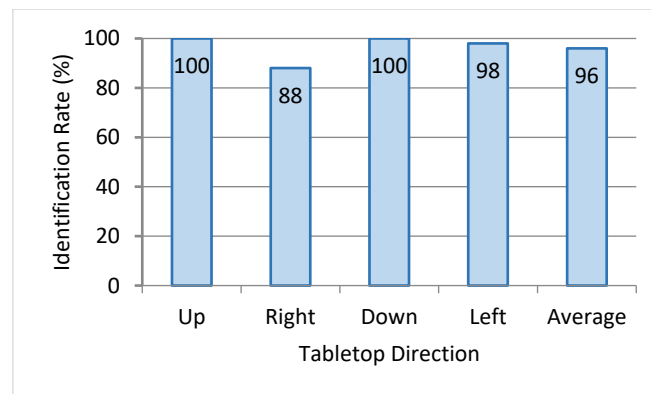


Figure 14. Identification rate for change-direction gesture.

In this regard, we note that the left/right edges of the table are shorter than the up/down edges. The experimental results in Figure 14 indicate that when users positioned in the left/right directions manipulate an object situated near an up/down edge, there is a tendency for the hand shadow to cross that up/down edge. Examples in which left/right hand direction could not be correctly identified are shown in Figure 15. These examples show users performing the change-direction gesture from the left and right directions. In either case, the number of pixels in the hand-shadow area intersecting the upper edge is greater than that intersecting the left or right edge. As a result, the system erroneously judges the operation to be that of the user positioned in the up direction.

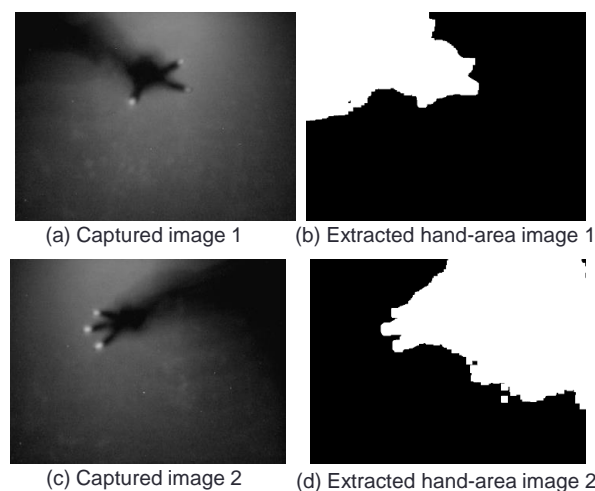


Figure 15. Change-direction gesture from left/right directions.

5.3. Recognition Accuracy Experiment for Copy Gesture

We performed a subject-based experiment for the copy gesture and measured its recognition rate. In this experiment, we divided eight male subjects in their 20s into two groups of four subjects each. With a subject positioned at each edge of the table, we had the subjects perform the copy gesture in two combinations—face-to-face across the table and side-by-side at neighboring edges—and measured the gesture recognition rate. Denoting the number of times the gesture was performed as N_{act} and the number of times the system recognized the copy gesture as N_{detect} , the recognition rate was computed by Eq. (11).

$$\text{Recognition rate} = \frac{N_{detect}}{N_{act}} \times 100 [\%] \quad (11)$$

The recognition rate for the copy gesture by eight subjects is shown in Figure 16. The overall average recognition rate of the copy gesture was approximately 85%. Gesture recognition could fail here if the direction of one of the two users performing this gesture could not be identified. An example of failing to identify the direction of one of two users is shown in Figure 17. In this case, two face-to-face users in the left and right directions are performing the copy gesture. Given a 2-finger touch gesture, this system would judge the operation to be a copy gesture provided that the user direction of each touch point could be identified and judged to be different. In the example of Figure 17, the position of the user on the right side of the image could be identified from the user's hand-shadow area. However, the head of the user on the left side creates a shadow, and as a result, the hand-shadow area cannot be distinguished from the dark portion of the background preventing the direction of that user from being identified.

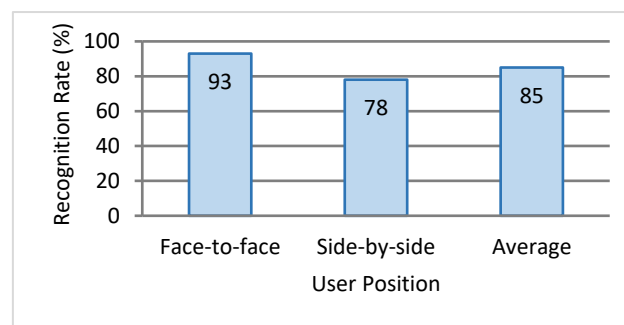


Figure 16. Recognition rate for copy gesture.



Figure 17. Example of failed recognition of a user's direction.

5.4. Results of System Usability Evaluation and Discussion

After asking male subjects in their 20s to freely use the system for about five minutes, we conducted a questionnaire-based survey using four subjects. This questionnaire consisted of questions based on the SUS evaluation method [17] and a section for free comments at the end. Ten questionnaire items based on the SUS evaluation method are given in Appendix.

In the survey, a subject responded to each questionnaire item on a five-point scale. In tabulating scores, we followed the SUS evaluation method that subtracts 1 from the score of each odd-numbered item and subtracts the score of each even-numbered item from 5. The average score of each item out of 4 points based on the SUS evaluation method is shown in Figure 18.

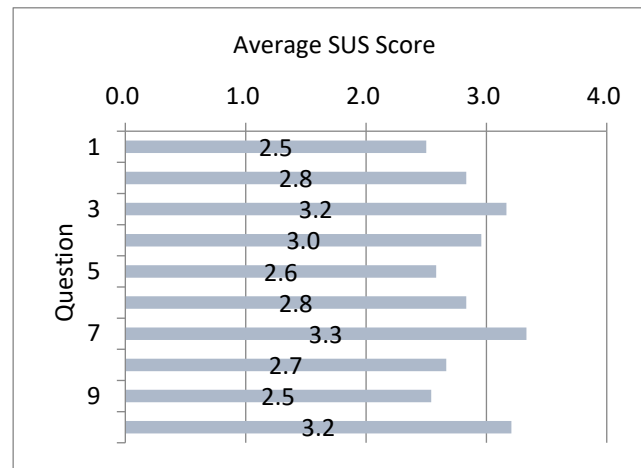


Figure 18. Average score for each item by SUS evaluation method.

In the SUS evaluation method, the scores for each questionnaire item are summed and then multiplied by 2.5 to convert to a 100-point scale. This value is taken to be the usability score. For the proposed system, the usability score by the SUS evaluation method was found to be 71.6 on average out of 100 points. In addition, a breakdown of questionnaire results revealed that items 3, 7, and 10 had high scores. These items are “3. I thought the system was easy to use,” “7. I would imagine that most people would learn to use this system very quickly,” and “10. There was no need to learn a lot of things before I could get going with this system.” Based on these results, the system was highly evaluated for making prior learning easy for the user and for being easy to operate.

In addition, the section for free comments included the statement “How about showing the users the results of user identification.” This opinion implies an operation of detecting a hand before it touches the table surface, identifying the position of that user, and visually notifying the user of that result with a cursor. We consider that presenting the users with system recognition results in near real time in this way should improve system operability. On the other hand, comments such as “Response of the move gesture is not so good” and “Trying to operate the system with two hands sometimes fails.” These comments reflect the need to improve system construction technology.

6. Conclusion

This paper described the development of a multi-touch tabletop system that identifies user position by infrared image recognition and presented the results of touch-gesture recognition accuracy experiments and a system usability evaluation. The proposed system picks up touch points and the shadow area of a user’s hand by infrared camera using an FTIR touch panel and infrared light and estimates the position of that user by image recognition. The multi-touch gestures prepared for this system include an operation to change the direction of an object and a copy operation in which two users generate duplicates of an object.

With this system, average recognition rate of the change-direction gesture and copy gesture were found to be 96% and 85%, respectively. The results of the questionnaire-based system usability evaluation, meanwhile, revealed that prior learning was easy for the user and that system operations could be easily performed. At the same time, opinions expressed in the free-comments section of the questionnaire indicated that further improvement in system construction technology was needed for advanced interaction.

Future research topics include enhancement of object-operation application functions and addition of an object deletion function, visual-support function by user-specific cursor display, and gesture functions that make use of user position identification. Going forward, an important goal will be to combine the research of tabletop systems with other technologies and to associate such systems with application fields in strong demand by society. For example, there is a high social need for combining tabletop systems with sensor technology in the field of medical treatment technology. In

addition, the problem of occlusion that arises when irradiating people with light is not limited to this system that uses infrared light. It must also be kept in mind that infrared light is affected by sunlight. To solve or prevent these problems, further studies are needed to achieve and apply technical advances.

Appendix

The System Usability Scale: When a SUS is used, participants are asked to score the following 10 items with one of five responses that range from Strongly Agree to Strongly Disagree:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

References

1. Taylor, J.; Bordeaux, L.; Cashman, T.; Corish, B.; Keskin, C.; Sharp, T.; Soto, E.; Sweeney, D.; Valentin, J.; Luff, B.; Topalian, A.; Wood, E.; Khamis, S.; Kohli, P.; Izadi, S.; Banks, R.; Fitzgibbon, A.; Shotton, J. Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences. *ACM Trans Graph* **2016**, *35*, 143:1–143:12, doi:10.1145/2897824.2925965.
2. Krupka, E.; Karmon, K.; Bloom, N.; Freedman, D.; Gurvich, I.; Hurvitz, A.; Leichter, I.; Smolin, Y.; Tzairi, Y.; Vinnikov, A.; Bar-Hillel, A. Toward realistic hands gesture interface: keeping it simple for developers and machines. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; CHI '17, Denver, CO, USA, 6-11 May 2017; ACM: New York, NY, USA, 2017; pp. 1887–1898.*
3. Shen, C.; Ryall, K.; Forlines, C.; Esenther, A.; Vernier, F. D.; Everitt, K.; Wu, M.; Wigdor, D.; Morris, M. R.; Hancock, M.; Tse, E. Informing the design of direct-touch tabletops. *IEEE Comput. Graph. Appl.* **2006**, *26*, 36–46, doi:10.1109/MCG.2006.109.
4. Morris, M. R.; Huang, A.; Paepcke, A.; Winograd, T. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI '06, Montréal, Québec, Canada, 22 - 27 April 2006; ACM: New York, NY, USA, 2006; pp. 1201–1210.*
5. Walker, G. A review of technologies for sensing contact location on the surface of a display: Review of touch technologies. *J. Soc. Inf. Disp.* **2012**, *20*, 413–440, doi:10.1002/jsid.100.
6. Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology; UIST '05, Seattle, Washington, USA, 23-26 October 2005; ACM: New York, NY, USA, 2005; pp. 115–118.*

7. Wobbrock, J. O.; Morris, M. R.; Wilson, A. D. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI '09*, Boston, MA, USA, 4–9 April 2009; ACM: New York, NY, USA, 2009; pp. 1083–1092.
8. Hinrichs, U.; Carpendale, S. Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI '11*, Vancouver, BC, Canada, 7–12 May 2011; ACM: New York, NY, USA, 2011; pp. 3023–3032.
9. Evans, A. C.; Wobbrock, J. O.; Davis, K. Modeling collaboration patterns on an interactive tabletop in a classroom setting. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; CSCW '16*, San Francisco, CA, USA, 27 February – 2 March 2016; ACM: New York, NY, USA, 2016; pp. 860–871.
10. Lundström, C.; Rydell, T.; Forsell, C.; Persson, A.; Ynnerman, A. Multi-touch table system for medical visualization: vpplication to orthopedic surgery planning. *IEEE Trans. Vis. Comput. Graph.* **2011**, *17*, 1775–1784, doi:10.1109/TVCG.2011.224.
11. Dietz, P.; Leigh, D. DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology; UIST '01*, Orlando, Florida, USA, 11–14 November 2001; ACM: New York, NY, USA, 2001; pp. 219–226.
12. Marquardt, N.; Kiemer, J.; Ledo, D.; Boring, S.; Greenberg, S. Designing user-, hand-, and handpart-aware tabletop interactions with the TouchID toolkit. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces; ITS '11*, Kobe, Japan, 13–16 November 2011; ACM: New York, NY, USA, 2011; pp. 21–30.
13. Lissermann, R.; Huber, J.; Schmitz, M.; Steimle, J.; Mühlhäuser, M. Permulin: mixed-focus collaboration on multi-view tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI '14*, Toronto, Canada, 28 April – 1 May 2014; ACM: New York, NY, USA, 2014; pp. 3191–3200.
14. Zhang, H.; Yang, X.-D.; Ens, B.; Liang, H.-N.; Boulanger, P.; Irani, P. See Me, See You: a lightweight method for discriminating user touches on tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI '12*, Austin, Texas, USA, 5–10 May 2012; ACM: New York, NY, USA, 2012; pp. 2327–2336.
15. Suto, S.; Shibusawa, S. A tabletop system using infrared image recognition for multi-user identification. In *Proceedings of Human-Computer Interaction – INTERACT 2013*; Cape Town, South Africa, 2–6 September 2013; Kotze, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. Eds.; Lecture Notes in Computer Science, 8118, Part II, Springer: Berlin, Heidelberg, 2013; pp. 55–62.
16. Suto S.; Shibusawa S. Evaluation of multi-user gesture on infrared-image-recognition based tabletop system. In *IPSJ Technical Report on the 157th Human Computer Interaction*; Tokyo, Japan, 13–14 March 2014; Information Processing Society of Japan: Tokyo, Japan, 2014; Vol. 2014-HCI-157, pp. 1–7.
17. System Usability Scale (SUS). Available online: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> (accessed on Feb 22, 2018).
18. Ohashi M.; Itou J.; Munemori J.; Matsushita M.; Matsuda M. Development and application of idea generation support system using table-top interface. *IPSJ J.* **2008**, *49*, 105–115.
19. Chen, L.; Day, T. W.; Tang, W.; John, N. W. Recent developments and future challenges in medical mixed reality. In *Proceedings of 2017 IEEE International Symposium on Mixed and*

- Augmented Reality*; ISMAR 2017, Nantes, France, 9-13 October 2017; IEEE Computer Society: Washington, DC, USA, 2017; pp. 123–135.
20. Genest, A. M.; Gutwin, C.; Tang, A.; Kalyn, M.; Ivkovic, Z. KinectArms: a toolkit for capturing and displaying arm embodiments in distributed tabletop groupware. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work; CSCW '13*, San Antonio, Texas, USA, 23–27 February 2013; ACM: New York, NY, USA, 2013; pp. 157–166.
 21. Zhang, Y.; Laput, G.; Harrison, C. Electrick: low-cost touch sensing using Electric field tomography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; CHI '17*, Denver, CO, USA, May 6-11, 2017; ACM: New York, NY, USA, 2017; pp. 1–14.
 22. North, C.; Dwyer, T.; Lee, B.; Fisher, D.; Isenberg, P.; Robertson, G.; Inkpen, K. Understanding multi-touch manipulation for surface computing. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II; INTERACT '09*, Uppsala, Sweden, 24-28 August 2009; Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R. O., Winckler, M., Eds.; Lecture Notes in Computer Science, 5727, Part II, Springer: Berlin, Heidelberg, 2009; pp. 236–249.
 23. Annett, M.; Grossman, T.; Wigdor, D.; Fitzmaurice, G. Medusa: a proximity-aware multi-touch tabletop. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology; UIST '11*, Santa Barbara, CA, USA, 16-19 October 2011; ACM: New York, NY, USA, 2011; pp. 337–346.