MOBAJES: Multi-user Gesture Interaction System with Wearable Mobile Device

Enkhbat Davaasuren and Jiro Tanaka

1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577 Japan
{enkhee,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. When people collaborate with multiple large screens, gesture interactions will be used widely. However, in conventional methods of gesture interaction, when there are multiple users, simultaneous interaction is difficult. In this study we have proposed a method using a wearable mobile device which enables multi-user and hand gestures only interactions. In our system, the user wears a camera-equipped mobile device like a pendant, and interacts with a large screen.

Keywords: Gesture, Gestural Interface, Large Screen, Mobile, Wearable Device, Multi-User.

1 Introduction

In the past years, large screen has been used more and more in various locations and situations, and their use will likely increase in the future. Many researchers have been performing research about large screen interaction methods. One of the most used interaction methods is Gesture Interaction, which is a method where the user can use body or hand gestures to interact with large screen. There are many types of gesture interaction systems, and each of them has good and weak points in multi-user interactions. In this research, we consider the hand gesture methods, and propose a hybrid interaction system that can work in a more stable manner in multi-user interactions (Fig. 1).

Fig. 1. System image

M. Kurosu (Ed.): Human-Computer Interaction, Part IV, HCII 2013, LNCS 8007, pp. 196–204, 2013.

1.1 Gesture Interface

Two kinds of gesture interfaces exist: wearable and non-wearable. In wearable interfaces, gestures are recognized by a sensor which is installed on the user's hand or body. In non-wearable interfaces, gestures are recognized by a camera which is attached to the large screen. The advantages and disadvantages of these two types are quite opposite. The advantage of non-wearable interfaces is that users do not need to wear any devices or markers, and this makes the system more mobile and easy to use. On the other hand, in a multi-user interaction, non-wearable interfaces are not so suitable. For example, problems such as calculation cost of gesture recognition or difficulty of identifying different users may occur. However, wearable interfaces also have disadvantages, like users needing to wear or hold in hand devices or markers. They are more suitable to multi-user interactions. This is because using a gesture recognizing device for each user makes it easier to identify the user, and the number of users will not affect the system calculation cost.

1.2 Proposal System

Our proposed system is a hybrid system, which applied both wearable and non-wearable advantages. We approached by wearing camera-equipped mobile device like a pendant as a gesture recognition device. In our system, user makes hand gestures in front of device which worn like a pendant, and make gesture interactions while seeing cursors of gestures on the large screen (Fig. 2). Since user does not need to wear device in hand, user can use both two hands freely, and can move freely even during the interaction. Also, user can concentrate to the interactions without thinking about device. Because each user has own device, user number will not affect to the system calculation cost.

Fig. 2. MOBAJES system

2 Related Work

In Gesture Pendant [1], researchers proposed new approach to detect hand gestures, but they did not consider multi-user. One important difference is that our system provides GUI feedback by using large screen during gesture interaction, while enabling more rich interactions to the user.

There are also many non-wearable gesture interaction systems such as [2], [3], [4], [5] and [6]; however, they still present multi-user interaction issues. In these systems, the whole recognition process is calculated in one place, and that makes the system unstable in case of multi-user interactions. We applied wearable approach to address these issues.

Systems about gesture interactions for public large screens ([7], [8], [9] and [10]) have also been developed. However, in these systems, the user needs to hold a mobile device in hand when interacting with large screen, which poses a burden to the user. In our system, user wears mobile device like a pendant, and thus does not burden the user.

The most related work to our system is Sixth Sense [11]. In the Sixth Sense system, user can display information on the other objects such as walls, by using a wearable projector, and make gestural interactions using markers of hand. In our system, the user can obtain more clear and rich information through the large screen, and can interact with bare hand gestures without markers.

3 MOBAJES Interactions

We developed a prototype system as a simple image manipulation system. In our prototype, user can manipulate the image files on the large screen using hand gestures, by wearing a mobile device like a pendant (Fig. 2). During the interaction, the user can see a cursor (Fig. 3 b) on the large screen as a feedback of gesture (Fig. 3 a). This will help users to understand each other's intention during multi-user interaction.

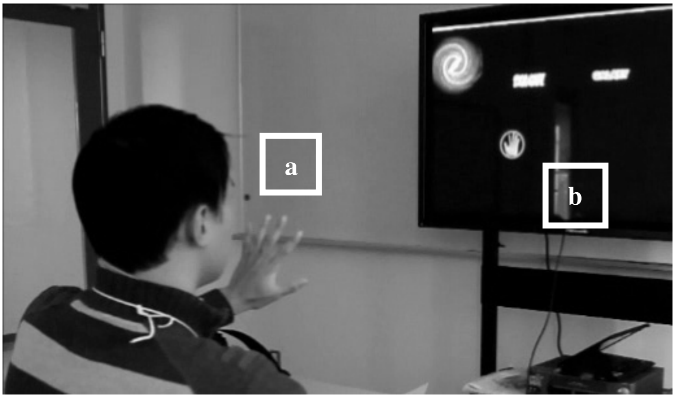

Fig. 3. Gesture feedback (a: user gesture, b: cursor on the large screen)

In this system, user can use 4 kinds of gesture such as “grab”, “release”, “point” and “L-letter” (Fig. 4). Those of the same shaped cursors will appear on the large screen.

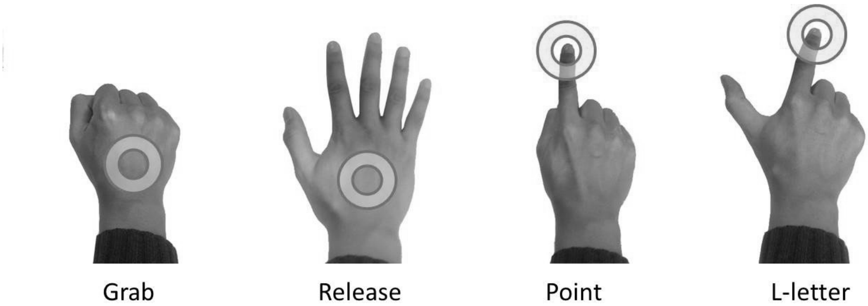

Fig. 4. Gesture types (Round markers represent the targeting point of each gestures)

Using these gestures, we implemented several basic interactions for our prototype, such as drag & drop, zoom & rotate and file share.

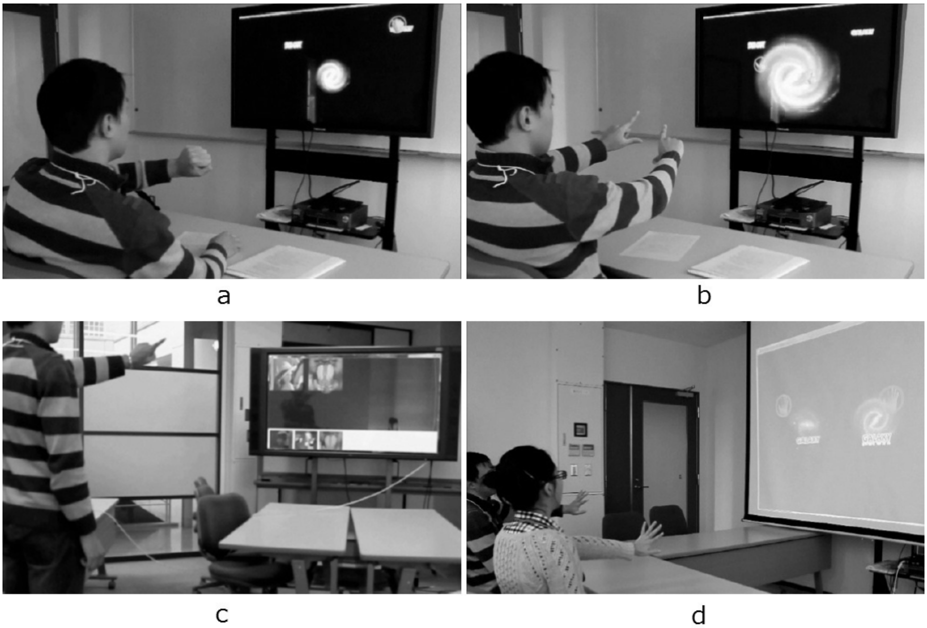

Fig. 5. Interactions (a: Drag & drop, b: Zoom & Rotate, c: File share, d: Multi-user interaction)

Drag & Drop. After user hovers over the target image file (using the cursor), he/she can drag the file with the “Grab” gesture, and can drop it with the “Release” gesture (Fig. 5-a).

Zoom & Rotate. User can zoom and rotate the file by using two hands (Fig. 5-b). To perform it, user needs to hover over the target image with tow hand's "Point" gesture, and while keeping that position, user needs to change the gesture to "L-letter". After that, user can zoom and rotate the file by changing the distance and the direction of two hands.

File Share. User can select the file on the screen using one-handed "Point" and "L-letter" gestures to copy it to his/her mobile device. To perform it, user need to hover over the target image file on the screen with "Point" gesture, and change the gesture to "L-letter". In reverse, user can display thumbnails of the image files of his/her mobile device on the large screen (Fig. 5-c), by changing gesture from "Release" to "L-letter". After displaying the thumbnails on the screen, user can copy and put the original file to large screen from mobile device by same gesture.

Multi-User Interaction. All these interactions can be performed in multi-user interactions as well (Fig. 5-d). Users can interact separately and collaboratively with each other, while knowing each other's intention.

4 Implementation

We have implemented our system using a camera-equipped Android mobile device, a large screen and Wi-Fi environment (Fig. 6).

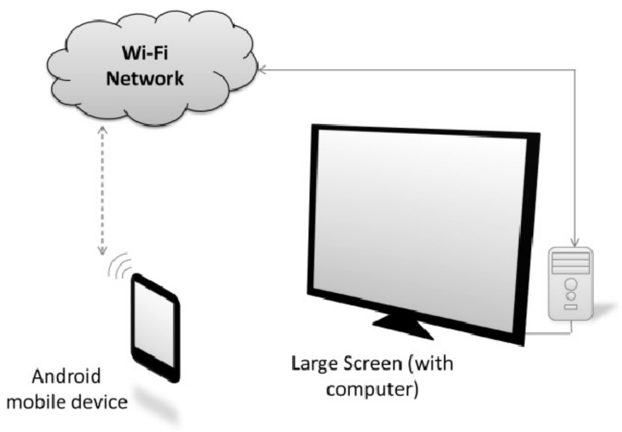

Fig. 6. System structure

4.1 Communications

Mobile devices and large screen will connect to one server application on the network, and communicate with each other using socket connection. The gesture information needs to be transferred in real time, so it is transferred by UDP protocol, between the mobile device and the large screen. Other information such as commands and files are transferred using the TCP protocol.

4.2 Gesture Recognition

The whole gesture recognition process is calculated by the mobile device. We used OpenCV[1] library and skin-color based method for recognizing gesture information. In order to recognize a hand gesture, we first detect skin-color areas (Fig. 7-b) from camera capture (Fig. 7-a). Using noise removal method, we can obtain clearer skin-color regions (Fig. 7-c).

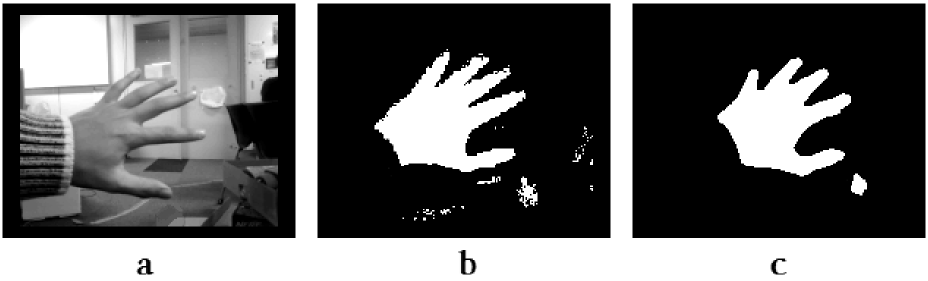

Fig. 7. Detecting skin color region (a: camera capture image, b: skin-color region with noise, c: clear skin-color region)

After that, we extract the contours of each skin-color regions (Fig. 8-a), and filter them by the area to obtain hand region contour (Fig. 8-b). Next, we extract the convex hull of hand region to detect finger-like parts. As can be noticed in the figure, finger tips belong both to the contour and the convex hull of the hand region (red parts in Fig. 8-c). We can use this fact to decrease the calculation cost and make recognition faster.

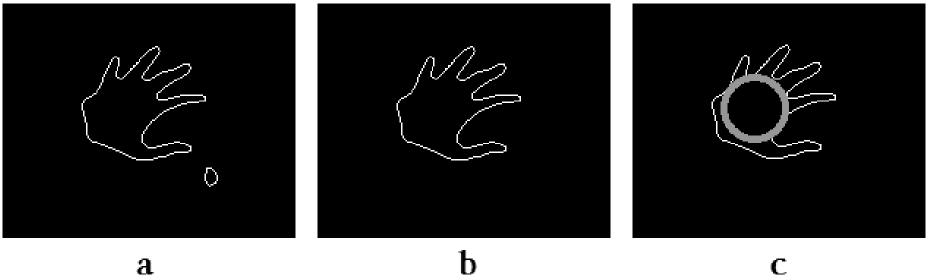

Fig. 8. Detection of hand region (a: contours of skin-color regions, b: contour of hand region, c: contour and convex hull of hand region)

[1] http://opencv.org

Next, we detect finger-like parts by using the angle of every tree points on the contour (Fig. 9-a). If the angle $\angle\theta$ is less than 30 degrees, it indicates the point P_i is the potential point of fingertip. Finally, we calculate the center of potential fingertip points as a real fingertip point (Fig. 9-b).

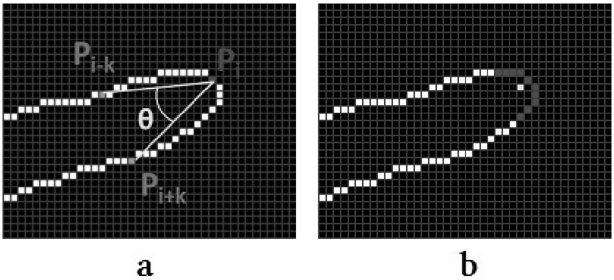

Fig. 9. Detection of fingertip (a: the angle between specific 3 points, b: found fingertip)

4.3 Noise Removal

To remove noise in gesture recognition process, we used the fact that the distance between user's hand and camera is almost constant (Fig. 10). If the distance is almost constant, we can assume the area size of hand region must be constant. By filtering the area size of all found regions, noises such as small size regions will be removed.

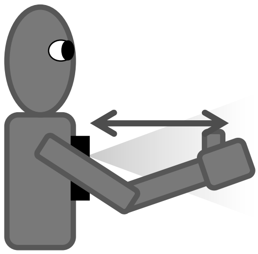

Fig. 10. Distance between user's hand and camera

We also used the natural fact that the hand region cannot be long and narrow shaped. To filter and remove long and narrow regions, we used the distance between the center point and the point closest to the center of each region. If the distance is less than specific threshold it means this shape is long and narrow and has to be removed.

After filtering the regions found by these limitations, we can obtain more clear hand regions for the further process.

5 Preliminary Evaluation

We performed a preliminary experiment to evaluate our system in multi-user interactions. In order to evaluate our system we asked two users to complete the given task in case of both single-user and multi-user interactions. After the experiment, we asked them about the difference between single-user and multi-user interactions. We also measured the amount of time needed to complete the task.

The task consisted of simply dragging and dropping a given picture to specified positions. To compare fairly, we used only one hand gesture for this task.

Our result shows that the performance of single-user interaction is almost same in both systems. And users said there is no difference between single-user interaction and multi-user interaction. Furthermore, in a multi-user interaction, knowing other user's intention by the cursor, it was easy to collaborate and avoid the collision.

In our next work, we will perform user study to know the error rates in each interactions. And to prove usefulness of our system, we will compare our system with a non-wearable gesture interface for large screen.

6 Summary and Future Work

In this study we proposed MOBAJES, a system which can intuitively interact with a large screen using a mobile device, and we implemented a prototype system. By wearing camera-equipped mobile device like a pendant, and performing hand gestures in front of the camera, the user can interact with large screen by gesture interaction. Since each user has own gesture recognition device, the gesture recognition cost does not affect the whole system cost, and user identification becomes very easy.

But, since we use skin-color detection method to recognize hand gestures, recognition accuracy easily affected by lighting of environment. We believe we need a more robust recognition algorithm suited to dynamic lighting change. In our future work, we will improve the recognition accuracy by implementing a more dynamic algorithm. We also intend to try a different device such as a depth sensor. Furthermore, we found that the gestures we use in our system can be tiring in case of a task taking a longer time. To address this problem, we need to consider easier gestures which users can perform easily and naturally without stress.

References

1. Gandy, M., Starner, T., Auxier, J., Ashbrook, D.: The Gesture Pendant: A Self-illuminating, Wearable. In: Infrared Computer Vision System for Home Automation Control and Medical Monitoring, ISWC 2000, pp. 87–94. IEEE Computer Society (2000)
2. Boulabiar, M.-I., Burger, T., Poirier, F., Coppin, G.: A low-cost natural user interaction based on a camera hand-gestures recognizer. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part II, HCII 2011. LNCS, vol. 6762, pp. 214–221. Springer, Heidelberg (2011)
3. Shi, J., Zhang, M., Pan, Z.: A real-time bimanual 3D interaction method based on bare-hand tracking. In: MM 2011, pp. 1073–1076. ACM (2011)

4. Bragdon, A., DeLine, R., Hinckley, K., Morris, M.R.: Code space: touch + air gesture hybrid interactions for supporting developer meetings. In: ITS 2011, pp. 212–221. ACM (2011)
5. Argyros, A.A., Lourakis, M.I.A.: Vision-based interpretation of Hand Gestures for Remote Control of a Computer Mouse. In: Huang, T.S., Sebe, N., Lew, M., Pavlović, V., Kölsch, M., Galata, A., Kisačanin, B. (eds.) HCI/ECCV 2006. LNCS, vol. 3979, pp. 40–51. Springer, Heidelberg (2006)
6. Clark, A., Dnser, A., Billinghurst, M., Piumsomboon, T., Altimira, D.: Seamless interaction in space. In: Proceedings of the 23rd Australian Computer-Human Interaction Conference (OzCHI 2011), pp. 88–97. ACM (2011)
7. Ballagas, R., Rohs, M., Sheridan, J.G.: Sweep and point and shoot: phonecam-based interactions for large public displays. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2005), pp. 1200–1203. ACM (2005)
8. Zhong, Y., Li, X., Fan, M., Shi, Y.: Doodle space: painting on a public display by cam-phone. In: Proceedings of the 2009 Workshop on Ambient Media Computing (AMC 2009), pp. 13–20. ACM (2009)
9. Jeon, S., Hwang, J., Kim, G.J., Billinghurst, M.: Interaction with large ubiquitous displays using camera-equipped mobile phones. Personal Ubiquitous Comput. 14(2), 83–94 (2010)
10. Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P.: Touch projector: mobile interaction through video. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI 2010), pp. 2287–2296. ACM (2010)
11. Mistry, P., Maes, P., Chang, L.: WUW - wear Ur world: a wearable gestural interface. In: Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA 2009), pp. 4111–4116. ACM (2009)