A Study of Followable User Interface to Hand Behavior

Takuya Yamaguchi, Kozo Mizutani, and Masayuki Arai

***Abstract*—In this work we study a followable user interface (FUI): a user interface that tracks the motion of the user's hand. In an FUI, the system provides a UI screen that tracks the motion of the user's palm. As a means of implementing such a system, we study a method involving an LCD projector and a depth sensor. In this paper we discuss methods for using the depth sensor to detect the position of the user's hand and for controlling the projection of the UI screen. Moreover, we discuss a method for detecting gestures in which the user places one hand on top of the other hand and makes sliding motions. As the result of constructing the prototype system with our methods, it is shown that the system was capable of properly projecting the UI to match the motion of the user's hand, to identify gestures, and to switch the UI based on this input.**

***Index Terms*—Hand tracking, projection mapping, depth sensor, followable user interface.**

I. Introduction

In an effort to identify one possible realization of next-generation ubiquitous computing technology, the authors have investigated methods for implementing a user interface (UI) that tracks the motion of the user's palm. The user interface is projected onto the palm of one hand of the user, who manipulates the UI through gestures with the other hand. By tracking the motion of the user's hand and adjusting the location of the projection accordingly, it is possible to provide the user with an experience that closely mimics the experience of holding a conventional smart device. The realization of such a UI would free users from the need to carry their smart devices with them and thus eliminate worries, such as dying batteries, present in conventional smart devices.

We refer to such a UI as a followable user interface (FUI). As one possible method of implementing an FUI, we have studied a scheme that uses liquid crystal display (LCD) projectors and depth sensors. In this paper we describe methods for determining the position of the user's hand and for controlling the position to which the UI is projected.

II. Background

Smart devices and wearable devices offer humans the capability of ubiquitous computing. In particular, the extreme portability of wearable devices is a significant expansion of the range of computing environments and their applications. Nonetheless, a number of difficulties remain to be overcome for next-generation ubiquitous computing; among these are the nuisance of having a device constantly affixed to one's body and the hassle of recharging batteries.

The FUI is a new type of UI that removes these drawbacks of conventional devices. Fig. 1 illustrates the concept of a FUI that uses projectors and sensors. The sensors positioned in the upper region of the environment capture the motion of the user. The system identifies the position of the user's hands from the captured data, then projects the UI to match that position. When the user moves within the environment, the position of the UI projection tracks this motion. Because the UI is displayed on the user's hand at all times, the system is able to provide an experience that closely mimics that of the user holding a device in his or her hand.

UIs that attempt to track the motion of a user's hand have been reported previously. The Lumipen system of Okumura *et al*. used a high-speed camera to capture hand motion and tracked this motion by using a mirror to control the projection axis of a projector [1], [2]. PALMbit [3], [4] and OmniTouch [5] considered a UI that tracked hand motion by using a projector and a depth sensor affixed to the user's body. In addition, the Hand-held Display System of Leung *et al*. dynamically projected the UI onto white cardboard [6], whereas the system of Lee *et al*. dynamically projected the UI onto a board equipped with markers in all four corners [7].

In the FUI considered here, projectors are installed at fixed locations. The actual UI is expressed on a portion of the projection area of the projectors. More specifically, the projection area of the projectors is the range that can be tracked. Because the projection area of the projectors is capable of expressing multiple UIs, the system can easily accommodate multiple simultaneous users. In addition, by considering methods that do not require boards or markers on which to project the UI, it may be possible for users to access the system without carrying any items.

In this regard, our proposal is similar to the LightSpace system of Wilson *et al.* [8]. LightSpace used a depth sensor and multiple projectors and implemented a framework in which objects were projected in a way that tracked the user's palm. However, although this system succeeded in realizing simple menus, the authors did not consider methods for allowing the user to use hand gestures to manipulate the projected UI.

III. Method

A. System Configurations for the FUI

The projectors and the sensors are mounted on the ceiling and directed downward. In this work we use a depth sensor to identify motions in the user's hand. The depth sensor

Manuscript received April 23, 2015; revised June 30, 2015.

Takuya Yamaguchi was with Department of Human Information Systems, Faculty of Science and Engineering, Teikyo University, Tochigi, Japan (e-mail: yamaguchi07t@gmail.com)

Kozo Mizutani and Masayuki Arai are with the Department of Information and Electronic Engineering, Facultly of Science and Engineering, Teikyo University, Tochigi, Japan (e-mail: mizutani@ics.teikyo-u.ac.jp, arai@ics.teikyo-u.ac.jp).

doi: 10.18178/ijke.2015.1.3.042

measures the distance from the sensor to a target object. The distance data Dt is defined as follows.

$$Dt = \begin{bmatrix} dt(0,0) & \cdots & dt(X,0) \\ \vdots & dt(x,y) & \vdots \\ dt(0,X) & \cdots & dt(X,Y) \end{bmatrix} \quad (1)$$

Here, t is the frame at a certain time of measurement, X is the width of the measurement range, Y is the height of the measurement range, and x and y are coordinates. Thus, the depth sensor allows the retrieval of distance data in a plane. Using this distance data allows us to identify the target object easily without influence from the brightness or color of its surroundings.

B. Hand Detection Using Depth Data

To detect the position of the user's hand when it is raised to the level of the user's chest, we use distance information obtained from the depth sensor to obtain two regions: the height of the head and the height of the chest. In this paper we refer to these regions as the head height area and the chest height area, respectively. The left panel of Fig. 2 shows the height ranges for these two regions. The right panel of Fig. 2 is a visualization of the distance information obtained from the depth sensor, with the head height area and the chest height area differentiated by color. We first determine the coordinates of the center of gravity of the head, G_h. Then we identify the coordinates of the point P_c that lies farthest from G_h within the chest height area. We take the point H — obtained by moving from P_c to G_h — as the coordinates of the user's palm.

C. Method for UI Projection Control to a User's Hand

The coordinates obtained for the user's hand are the coordinate system for the depth sensor. Processing is required to reconcile the projector coordinate system with the depth sensor coordinate system and to project the UI onto the user's palm. For this purpose, we adopt a technique for adjusting the detection area of the depth sensor based on the projection area of one of the projectors. Fig. 3 shows a screenshot captured from the program that specifies the detection area of the depth sensor. First, the projector projects onto a white screen. In addition to capturing this situation with the depth sensor, we specify the detection area by choosing the selection border to match the projection area.

Fig. 4 schematically illustrates the detection area and the actual projector output. As noted above, the detection area is specified by matching the projection area. At the position detected for the user's hand within the detection area, an object representing the UI is displayed. By filling in all other regions with black, we ensure that the UI is projected only onto the user's palm.

D. Gesture Detection

In anticipation of a user's attempts to manipulate the projected UI, we studied techniques for detecting gestures in which the user places one hand on top of the other hand and makes sliding motions. These gestures allow the realization of operations such as switching the entire UI or scrolling it from left to right.

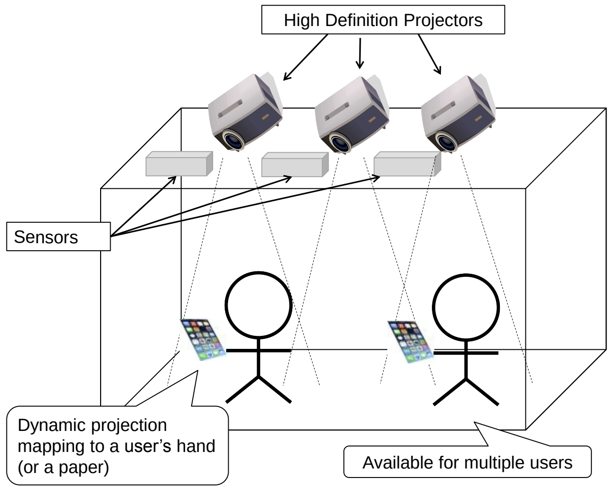

Fig. 1. Concept of a followable user interface.

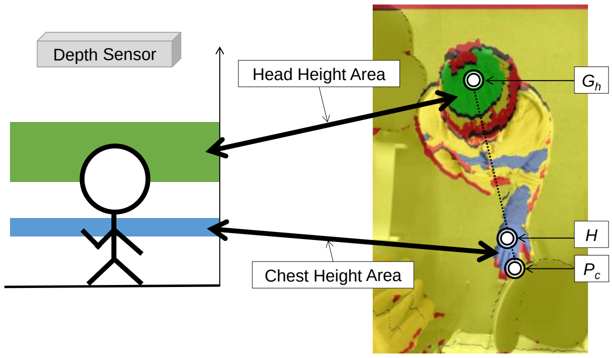

Fig. 2. Hand detection using depth data.

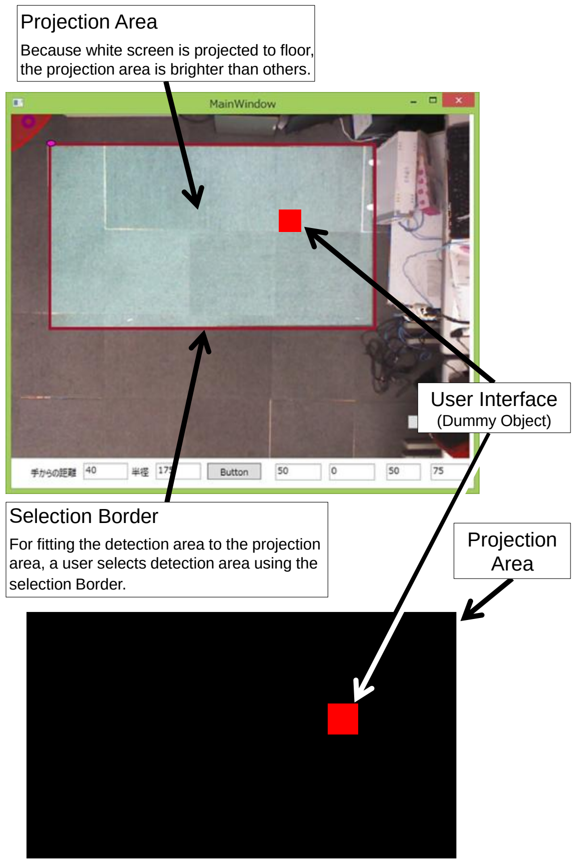

Fig. 3. Program to select the detection area of the depth sensor.

Fig. 5 illustrates the concept of gesture detection processing. First, we designate the area outside a certain range surrounding G_h (the area outside the black mask in Fig. 5) as the gesture area. Before the gesture begins, the detected position of hand H roughly agrees with the center of gravity of the chest height area G_c in the gesture area. When the gesture begins, G_c moves in the direction of the moving hand. In addition, the chest height area (S_c) varies from the area of a single hand to the area of both hands. When the two hands are fully separated, area S_c becomes the area of both hands and ceases to change. At this point we detect that the gesture is complete, and we take the direction in which G_c has moved relative to H as the direction of the gesture.

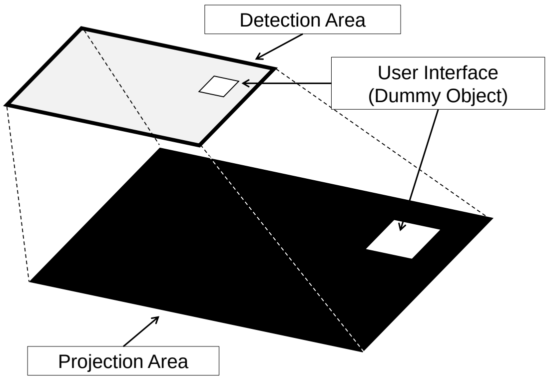

Fig. 4. Concept of mapping the detection area to the projection area.

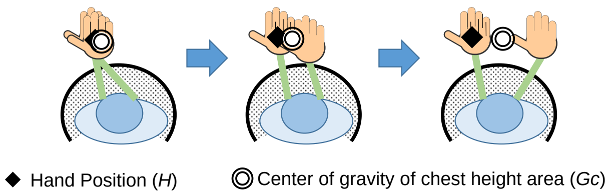

Fig. 5. Concept of gesture detection method for sliding hands.

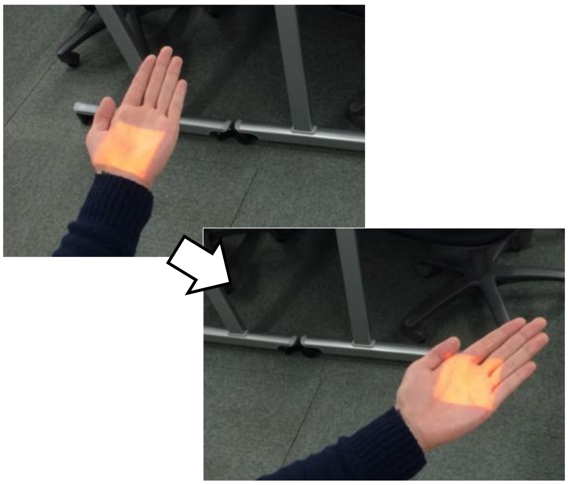

Fig. 6. Example of the FUI by the prototype system.

IV. Experiment

We tested the methods discussed in the previous section by constructing an actual prototype system. For the depth sensor we used a Kinect device. Table I lists other items used to construct the prototype, as well as various experimental conditions, such as the projector model and the positions at which the sensor was installed.

Fig. 6 depicts an example of our prototype system in operation. To simplify the detection step, the UI display consists of only a colored rectangular box. We implemented a procedure for changing the color of the box when a gesture is detected (Fig. 7).

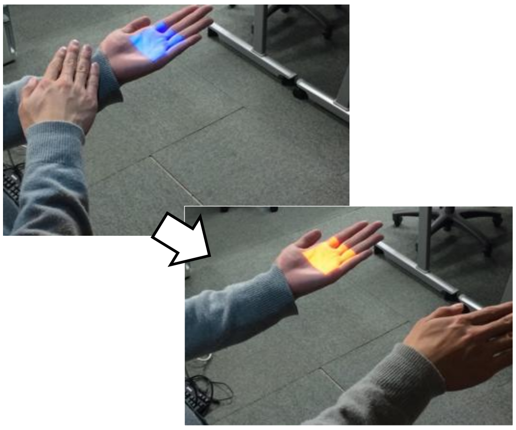

Fig. 7. Example of gesture detection, shown by the color change.

TABLE I: Development Tools and Experimental Environment

Development Tools for the Prototype System		
	IDE	Microsoft Visual Studio 2013
	API	Kinect for Windows SDK v1.8
Experimental Equipment		
PC	OS	Windows 8.1
	CPU	Core-i7-4790
	RAM	8 GB
Projector	Model	EPSON
	Resolution	1280 × 800
Depth Sensor	Model	Kinect v1
System Configuration		
Projector	Height from Floor	2400 mm
Depth Sensor	Height from Floor	2400 mm
Head Area	Height from Floor	1600–1900 mm
Hand Area (Chest area)	Height from Floor	1100–1200 mm

V. Consideration

We verified that our proposed method successfully tracks the user's hand motions and projects the UI accordingly. We confirmed that our system is able to track hand motions not only in the left-right direction but also in the front-back direction. We also verified that our system can detect gestures in which the user puts both hands together, then makes sliding motions with one hand. The system is able to properly identify the direction of the sliding motions. The sliding gestures switch the content of the UI.

At present, we have specified fixed values for the head height area and the chest height area. To allow our system to be used by individuals of different heights, the system will eventually require a framework for dynamically determining these values. In addition, the method discussed here is not well suited for detecting hand positions in cases involving multiple simultaneous users. Accommodating such situations will require the study of new detection methods. For example, when the position of a hand lies outside a certain fixed range based on the value of G_h, we might identify the hand as belonging to a different user.

In addition, the realization of a practical UI will require detection not only of motions involving the entire palm of the hand, but also gestures involving only motions of the fingertips. For this purpose it will be essential to improve the accuracy of our detection methods by introducing additional types of sensors.

VI. Conclusion

In this paper we investigated methods using an LCD projector and a depth sensor to implement an FUI that tracks the palm of the user's hand. We constructed an actual prototype of our system and evaluated its performance. We demonstrated that the system was capable of properly projecting the UI to match the motion of the user's hand, to identify gestures, and to switch the UI based on this input. Future work will address methods for achieving practical realization of this system.

References

[1] K. Okumura, H. Oku, and M. Ishikawa, "Lumipen: Projection-based mixed reality for dynamic objects," in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, pp. 699-704, 2012.

[2] K. Okumura, H. Oku, and M. Ishikawa, "Acitve projection are using high-speed optical axis control and appearance estimation algorithm," in *Proc. 2013 IEEE International Conference on Multimedia and Expo*, 2013, pp. 1-6.

[3] G. Yamamoto and K. Sato, "PALMbit: A body interface utilizing light projection onto palms," *Jour. of the Inst. of Image Information and Television Engineers*, vol. 61, no. 6, pp. 797-804, 2007.

[4] G. Yamamoto, H. Xu, K. Ikeda, and K. Sato, "PALMbit-Silhouette: A user interface by superimposing palm-silhouette to access wall displays," *Human-Computer Interaction. Novel Interaction Methods and Techniques Lecture Notes in Computer Science*, vol. 5611, pp. 281-290, 2009.

[5] C. Harrison, H. Benko, and A. D. Wilson, "OmniTouch: Wearable multitouch interaction everywhere," in *Proc. the 24th Annual ACM Symposium on User Interface Software and Technology*, 2011, pp. 441-450.

[6] M. C. Leung, K. K. Lee, K. H. Wong, and M. M. Y. Chang, "A projector-based movable hand-held display system," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, 2009, pp. 20-25.

[7] J. C. Lee, S. E. Hudson, J. W. Summet, and P. H. Dietz, "Moveable interactive projected displays using projector based tracking," in *Proc. the 18th Annual ACM Symposium on User Interface Software and Technology (UIST 05)*, 2005, pp. 63-72.

[8] A. D. Wilson and H. Benko, "Combining multiple depth cameras and projectors for interactions on, above and between surfaces," in *Proc. the 23nd Annual ACM Symposium on User Interface Software and Technology*, 2010, pp. 273-282.

Takuya Yamaguchi graduated from the Department of Human Information System, Faculty of Science and Engineering, Teikyo University, Japan in 2015. He mainly engaged in information processing. At present, he works in Meiji Yasuda System Technology Company Limited.

Kozo Mizutani received B.E. degree from Teikyo University in 2000. He worked Fujisoft ABC Inc. (now Fujisoft Inc.) in 2000 and BB Technology Corp. (now SoftBank BB Corp.) in 2002. He received M.E. and D.E. degrees from Graduate School of Science and Engineering Teikyo University, in 2005 and 2008. He worked as an assistant professor at Medical Information and Systems Research Center, Teikyo University from 2008. From 2014, he has been working as a Lecturer at Department of Information and Electronic Engineering, Faculty of Science and Engineering, Teikyo University. He is a member of IEICE, IPSJ, JSiSE and JAMI.

Masayuki Arai is a professor in the Graduate School of Sciences and Engineering at Teikyo University. He received his B.E. degree from Tokyo University of Science in 1981 and Dr. Eng. degree from Utsunomiya University in 1995. His research interests include pattern recognition, natural language processing and information visualization. He is a member of the Information Processing Society of Japan and IEEE.