Pattern-based Gesture Interaction with a Digital Table

Gunhee Kim, Hyunchul Cho, Wen-hua Pei, Ji-Hyung Park, and Sungdo Ha
Korea Institute of Science and Technology
39-1 Hawolgok-Dong, Wolsong-Gil 5, Seongbuk-Gu, Seoul, Korea
{kani, hccho, peach, jhpark, s.ha}@kist.re.kr

Abstract

Various applications of tabletop surface system have been developed recently. By merging dis play and manipulating space, we can take natural gestures as input in these applications. However, their interaction methods do not much differ much from traditional methods despite the change of user interface. Users still encounter a cognitive load in that they should know the defined mapping relations between gestures and functions.

In this paper, we design a system framework for users to interact naturally as if manipulating real entities on a table. In this framework, users can reduce their burden of awareness regarding interaction rules, and the system can understand users' gestures with a knowledge model that presents three elements: gestures, features and patterns.

1. Introduction

Recently, various digital tabletop applications have been developed for more natural and collaborative works. An important merits of these applications is that users can see and touch digital contents directly because the digital table can be both the display and manipulating space. Such an interface can make human-computer interaction more flexible and natural. Gesture interaction is a particularly ideal form which makes intuitive manipulation possible.

Previous studies on gesture interaction have focused on how to design gestures so they are simpler and easier to use. However, their interaction methods are not much different from desktop applications. These applications have the same interaction structure in that users should know the mapping relations between gestures and functions. For easy and natural human-computer interaction, users should be able to do user-centered gestures which are used in the real world, and the system should be able to respond to these gestures understanding users' intention.

Our objective of this research is to develop a framework that can enable users to interact naturally without an additional burden to aware interaction rules. Also, to understand various gestures, we designed a knowledge model that presents features and patterns of gesture. As a result, the system is able to infer how to respond to users' gestures. To illustrate the validation of this framework and knowledge model, we implement the system prototype which includes example gestures within a digital tabletop [figure 1].

2. Related Works

Early research on gesture interactions has mainly concentrated on how to recognize users' gestures. However, with development of direct-touch surface technology, it has been a critical i ssue how to map naturally between gestural input and system functions. Wu and Balakrishnan [1] designed multi-finger and whole hand gestures categorized by the shape of hand contact with surface on the DiamondTouch [2]. A study with more specific targets by Apted et al. [3] presented gestures associated with sharing digital photos for the elderly and evaluated them using task scenarios. Hilliges et al. [4] also described interactions for browsing, sorting and sharing digital photo collections. For more effective interactions, Tse et al. [5] combined speech and gesture together. They showed these multimodal commands can simplify users' tasks into interacting with digital tabletops.

Also, studies on methodology concerning gesture design and interaction have been undertaken. Wu et al. [6] presented three design principles: gesture registration, relaxation, and reuse. Combining these principles, they illustrated that

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

compound gestures created from a sequence of simple gestures which could be unique or reused. K. Everitt et al. [7] presented Modal Spaces that can enhance interfaces to respond differently to the same gestures depending on where users execute them.

3. Framework

We present a framework which enables users to interact naturally with hand gestures on a digital table. In general tabletop applications, one gesture is mapped to one function or command. Therefore, users should know these relations, and make predefined gestures as input. In contrast, users can make input gesture without cognitive load in our system. Instead of burdening users, the system possesses knowledge about gesture interaction, and infers proactively users' gestures and intentions. When users make a gesture on the digital surface, our system begins to analyze the gestures and designs the response according to users' intention. In order to understand users' gestures, the system performs three analysis steps [figure 2]. First, when users make gestures as input, the system extracts their features, and then finds proper interaction patterns with them. Finally, it designs how to respond to gestural input by the objectives and properties described in the gesture patterns.

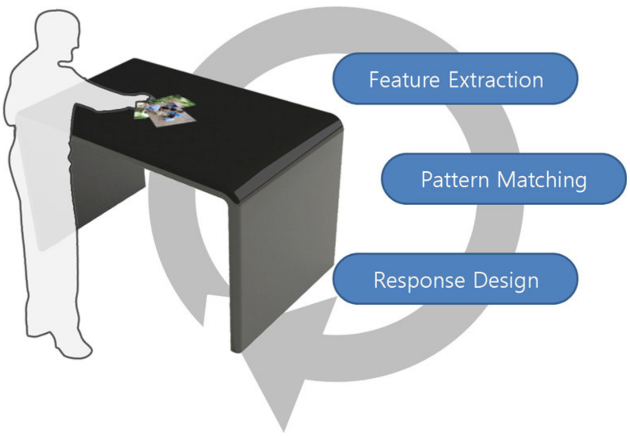


Figure 2: System framework

3.1. Feature extraction

Feature extraction is the first step to understand users' gestures. Because the form of gestural input can be expressed variously, the gestural interaction system should catch the features of related interactions. These features are used for interpreting gesture meaning. For example, if a user sweeps the surface with a side of a hand, the system can extract the following features: selection by outside of contents, dragging movements, and one direction [figure 3].

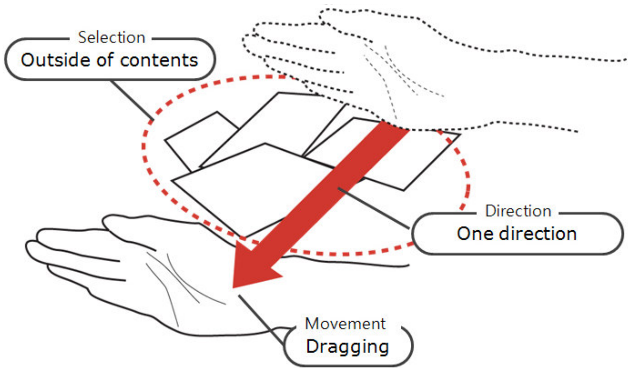


Figure 3. An example of feature extraction

3.2. Pattern matching

After extracting features, the system finds proper interaction patterns with matching rules between gestural features and interaction patterns. For example, the interaction pattern, 'Moving all contents with a sweeping motion' can be found from the previously illustrated features: selection by outside of contents, movement by drag, and same direction of touched points.

3.3. Response design

The goal and properties of interaction are described in interaction patterns. Thus, our system can design responses based on these descriptions. Although gestural inputs are similar, interaction results can be different unless patterns are the same. Figure 4 presents the difference between two interactions. Only initially selected contents are moved in the left figure, but in the right, all contents in the area are moved. These results are made possible by goal and property differences described in the patterns.

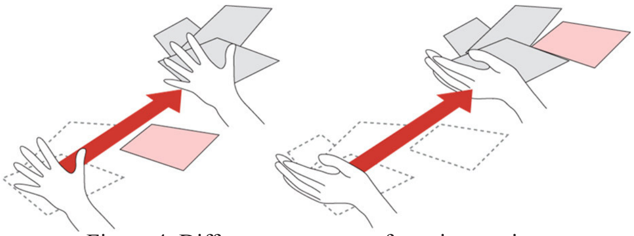


Figure 4. Different responses of two interactions

4. Knowledge Model

In this chapter, we present a knowledge model for gesture interactions. For understanding users' gestures, the system should have interaction knowledge about gestures and their meanings. In order to build a more detailed model, we have defined the use cases, manipulating digital photographs on the digital table. When handling digital photos, people make various gestures as input to perform interaction functions such as moving, rotating and resizing.

First, we have collected and analyzed users' gestures to design gesture features and patterns. We observed user actions when people manipulated photos on a table. Eighteen people participated in our user study, and they performed the predesigned fourteen tasks with hand gestures. For collecting more intuitive and natural gestures, the experimental environment was designed with a real table and photos. People, first, listened to explanation about the task which they were going to do, and then acted without any constraints. While they performed tasks, we observed their gestural actions: selecting, moving, rotating, resizing, deleting, restoring, and aligning photos. After observing their gestures, we classified observed gestures into typical forms. These results were used to design features and patterns.

4.1. Features

We have defined gesture features from observed data. These features mean essential factors which influence gestural interactions. In the system, they are used to distinguish users' gestures and to find interaction patterns. When subjects are moving, rotating, and resizing contents, we have designed seven features with three criteria: selection, movement, and direction. The method of target selection has two types: touching on contents directly and setting the range by outer position [figure 5(a)]. The movement type is also divided into two: drag movements which assign both the initial and the last position, and slide movements which deliver force to targets at the initial position [figure 5(b)]. The combinations of directions and distances between touched points are also defined as the feature [figure 5(c)].

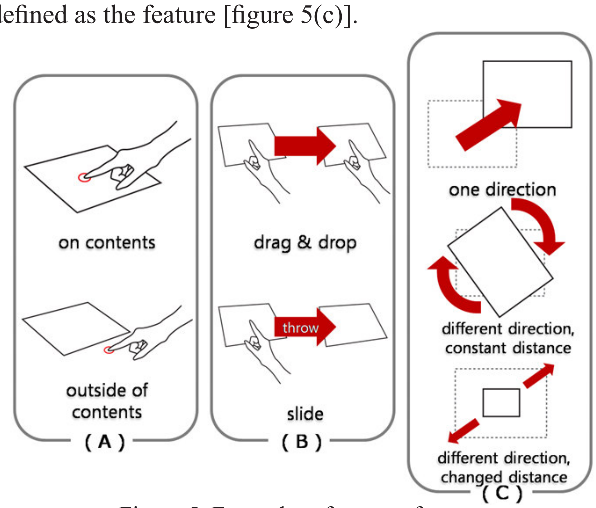


Figure 5. Examples of gesture features

4.2. Patterns

Interactions patterns are a means to express the meaning of gestural input in a standard form. These patterns describes the objective which reflects users' intentions, and also has properties which can be guidelines as to how to respond to users' gestures. The patterns are determined by a pattern matching process, and these are used for designing responses. For interaction to move, rotate, and resize contents, we have designed eighteen interaction patterns. Figure 6 shows an example of the pattern, 'Moving all contents with a sweeping motion'. The goal of this pattern is to move all content touched by a palm. Generally, people make this sweeping gesture when they would like to move contents of a wide range at one time. All contents in the area where a user's a hand passed by are moved without exception even if the contents are sand or water drops. After making this gesture, contents are gathered nearby the hand although their initial positions are apart from each other.

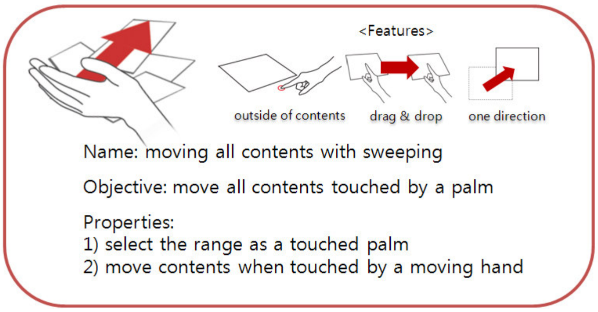


Figure 6. An example of an interaction pattern

5. Implementation

We now describe a system prototype to validate our proposed frame. Users can manipulate photos with this application.

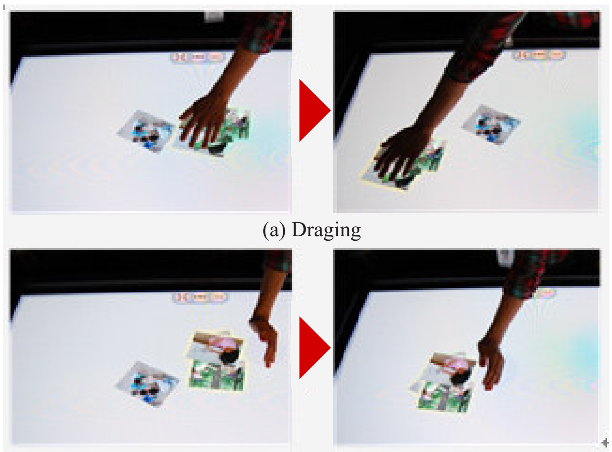
5.1. System hardware

We used the tabletop surface system which consisted of a PDP-based table display and IR sensing module. The table display consists of a 50 inch PDP with XVGA (1280x768) resolution for the visual display and a table frame to support the PDP. Unlike the projection-based display, the PDP-based display does not require dark lighting condition and calibration. In order to track the absolute position of a hand, a commercial infrared imaging sensor (XYFer from E-IT) [8] is installed on the table display. It can sense two touches on the display at the same time quickly and accurately.

5.2. Prototype application

We developed a digital tabletop system for intelligent and responsive space [9]. In the environment, a prototype application has been designed for manipulating photos by gestural input. The application supports that users can take various gestures as input to perform several functions: selecting, moving, rotating, and resizing. When users make gestures, the application analyzes and responds to users' intentions with the knowledge model.

Figure 7 shows a user making gestures on our tabletop surface. In the upper picture, the user touches photos and drags them. From this gesture, only the initially selected contents are moved. In the lower picture, by contrast, we can see all photos are moved following the hand.



(a) Draging

(b) Sweeping

Figure 7: Moving photos with different gestures

6. Conclusion

We have presented a framework that allows one to understand users' various gestures with a knowledge model. In this tabletop environment, users can interact naturally with barehanded gestures like when manipulating contents on a real table. In addition, we have designed a gesture knowledge model which includes gesture features, and interaction patterns. It is needed for the system to recognize and to respond to users' gestures. This approach to understanding users' natural gestures can be particularly useful to people who are not skillful with digital devices. Actually, they are afraid of using applications because of their unfamiliar interfaces.

In the future, we plan to extend our framework that is only concerned with the context of gestures as of now. For more intelligent interaction, it is necessary to evaluate the task content in the context of each particular environment.

References

[1] M. Wu and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. UIST'03, pp.193-202, 2003.

[2] P. Dietz and D. Leigh. DiamondTouch: A multi-user touch technology. UIST'01, pp.219-226, 2001.

[3] T. Apted, J. Kay, and A. Quigley. Tabletop sharing of digital photographs for the elderly. CHI'06. pp. 781-790, 2006.

[4] O. Hilliges, D. Baur, and A. Butz. Photohelix: Browsing, sorting and sharing digital photo collections. TABLETOP'07, pp.87-94, 2007.

[5] E. Tse, C. Shen, S. Greenberg, and C. Forlines. Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. Proc. Working Conf. advanced visual interfaces (AVI), pp.336-343, 2006.

[6] Wu, M., & C. Shen, K. Ryall, C. Forlines, and R. Balakrishnan. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. TABLETOP'06, pp.183-190, 2006.

[7] K. Everitt C Shen, K Ryall, and C Forlines. Modal Spaces: Spatial multiplexing to mediate direct-touch input on large displays. conference on human factors in computing systems, pp.1359-1362, 2005.

[8] XYFer system, http://www.e-it.co.jp

[9] J. Park, K. Yeom, and J. Lee. Intelligent and Responsive Spaces: integrated design of ambient physical spaces and information spaces. Communications of the Korea Information Science Society, 26(3):15-25, 2008