MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Modal Spaces: Spatial Multiplexing to Mediate Direct-Touch Input on Large Displays

Katherine Everitt, Chia Shen, Kathy Ryall, Clifton Forlines

TR2005-002 April 2005

Abstract

We present a new interaction technique for large direct-touch displays called Modal Spaces. Modal interfaces require the user to keep track of the state of the system. The Modal Spaces technique adds screen location aas an additional parameter of the interaction. Each modal region on the display supports a particular set of input actions and the visual background indicates the spaceś use. This ẅorkbech approachëxploits the larger form factor of display. Our spatial multiplexing of the display supports a document-centric paradigm (as opposed to application-centric), enabling input gesture reuse, while complementing and enhancing the current existing practices of modal interfaces. We present a proof-of-concept system and discuss porential applications, design issues, and future research directions.

CHI 2005

Modal Spaces: Spatial Multiplexing to Mediate Direct-Touch Input on Large Displays

Katherine Everitt[1,2] **Chia Shen**[1] **Kathy Ryall**[1] **Clifton Forlines**[1]

[1]Mitsubishi Electric Research Labs
201 Broadway
Cambridge, MA, 02139, USA
{shen, ryall, forlines}@merl.com

[2]DUB Group
Department of Computer Science and Engineering
University of Washington, Seattle, WA, USA
everitt@cs.washington.edu

Abstract

We present a new interaction technique for large direct-touch displays called Modal Spaces. Modal interfaces require the user to keep track of the state of the system. The Modal Spaces technique adds screen location as an additional parameter of the interaction. Each modal region on the display supports a particular set of input actions and the visual background indicates the space's use. This "workbench approach" exploits the larger form factor of display. Our spatial multiplexing of the display supports a document-centric paradigm (as opposed to application-centric), enabling input gesture reuse, while complementing and enhancing the current existing practices of modal interfaces. We present a proof-of-concept system and discuss potential applications, design issues, and future research directions.

Categories & Subject Descriptors: H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms: Design, Human Factors.

Keywords: Modal, Gesture, Large Screen Displays, Tabletop Interaction, Direct-touch input.

INTRODUCTION

Modal interfaces serve a useful role in many applications by providing users access to a large number of commands through a much smaller set of input actions. For a modal interface to be successful it is critical that users know what mode they are currently in. Visual feedback is most often used to indicate mode. For mouse-based interaction, one current paradigm is to change the on-screen mouse pointer to reflect the current mode. Thus the pointer serves two roles – it provides positional as well as state information to the user. This same solution, however, does not work well with large direct-touch interaction displays such as interactive tabletops and electronic whiteboards. Another feedback option is shadowing a toolbar button, but toolbars are not always reachable in large displays and it can be difficult to share toolbar feedback between multiple people.

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

Key differences in the visibility of the feedback information available to the user between direct-touch interfaces (such as tabletops and whiteboards) and relative input mouse-based interfaces are (1) the on-screen pointer is not present when the input device (e.g., hand or stylus) is out of touch sensing range, (2) the tracking of the pointer with the inputting device often trails inaccurately, and (3) the arm or hand frequently occludes the pointer.

On the other hand, direct-touch large displays can potentially support a broader set of user input actions than a mouse-based system. Touch-based multi-finger gestures can be used to express a complex set of operations in computational systems. Although a large gesture set may be hard to learn and remember, and reliable gesture recognition for large gesture sets is still an open research question, the recognition and usage of a small set of multi-finger touch gestures has already been shown to be feasible [4]. In addition, if we can support gesture reuse, i.e., using the same gesture under different mode for different tasks, people will be able to work with a small set of gestures while accessing a large set of commands.

In this paper, we introduce Modal Spaces, a "workbench approach" to enhance support for modal interfaces on large, direct-touch input displays. On a physical workbench, location and tool appearance play a key role; there are different spaces used for different tasks. For example, one space is for soldering connections, another for storing wires, and yet another for cutting them. It is always clear which actions occur at which specific location. Similarly, with Modal Spaces, we explore how to effectively combine a large screen with multiple input actions. Taking a document-centric view, mode is changed for a document when it is physically moved from one modal space to another. We use spatial multiplexing to support direct-touch input in a way that traditional mouse-based modal interfaces do not.

In this paper we discuss related work and then present Modal Spaces through a simple proof-of-concept image manipulation application. We use spatial multiplexing on the display to mediate direct-touch inputs; adjusting the meaning of three gestures. By splitting up the workspace into modal regions and by reusing the same gestures across the different modal regions, the usefulness of a small gesture set is then enhanced. Thus, a user can focus on the task or operation to be performed, rather than current mode.

RELATED WORK

Modal Spaces are not without precedent in interface design. For example, the PalmPilot PDA has separate regions for performing strokes that are interpreted as numbers and strokes that are interpreted as letters. This simple segmentation of the input space reduced input errors without requiring the user to learn a more complex (i.e. more easily distinguishable via software) set of gestures. While the PDA example maps strokes by regions, we add the additional parameter of object type, and consider a more sophisticated usage model.

Toolglass widgets [1] are user interface tools that can appear, as though on a transparent sheet of glass, between an application and a traditional pointer. They can be positioned with one hand while the other hand positions the pointer. They provide enhanced tool capability while still being very visible. They can even be layered on top of each other for additive functionality. Modal Spaces share some of these capabilities. Although Modal Spaces cannot be layered, Modal Spaces can change how an object is displayed. Additionally, a set of Modal Spaces form a conceptual space where a set of objects can be operated on as a group. Toolglasses are complimentary to, and can be used in conjunction with, Modal Spaces.

Some notion of spatial multiplexing also exists in a conventional GUI. Different windows on a computer sometimes deal with mouse actions differently, in ways that map to their functionality. For example, in a paint program, a left click maps to distinct actions based on location: drawing on the canvas, selection over a tool palate and resizing over the application's title bar. People quickly learn the different effects that their mouse has in different regions of the interface. In contrast, Modal Spaces uses a document-centric view, so what matters is the location of the document. Actions occur on documents rather than in application toolbars or backgrounds.

MODAL SPACES

There are challenges involved in using traditional methods for dealing with the mode problem on large displays. Menus and toolbars may not be reachable from all sides. Also, when there are multiple people using the interface, as is often the case with large displays, then there can be confusion as to who is in what mode, or what mode the group is in.

The Modal Spaces technique is an enhancement to conventional modal interfaces. It addresses two critical issues for direct-touch input large displays: screen-location-based indication of the current mode, and provision for a seamless method for changing between modes. It replaces mode with location in the regular object+action+mode equation for commands, instead commands are based on object+action+location. With Modal Spaces, the display is divided into multiple semantic workspaces, called modal regions. Each region represents a different mode, and maps users' actions to a different set of system commands. Thus location plays a role in mediating user input. A particular object or document will respond differently to the same action depending on the modal region in which it is positioned. For example, a touch in one modal region could open a popup menu while the same touch in another modal region could begin a stroke operation. We are mapping different interaction rules to different spaces.

Proof-of-Concept

In this section, we present a simple example of a Modal Spaces system for image-modification applications. This proof-of-concept system has been designed and implemented on a DiamondTouch [2] direct-touch interactive tabletop. There are four different regions in our design: The Layout Region, Cutting Region, Resizing Region, and Annotation Region. Since DiamondTouch devices can sense multiple touch points and areas, our example Modal Spaces system supports three touch interaction primitives: touching the table with one, two, or three fingers at a time. Our example is meant to illustrate the Modal Spaces concept; the issues surrounding the choice of gesture set, while important, are beyond the scope of this paper.

Figure 1 shows the Modal Spaces in our tabletop system, and Table 1 illustrates the mapping of the three interaction primitives to the four modal regions. Seven actions are available using only three gestures. In the Layout Region, the user can move a document with any touch and drag action using one, two or three fingers. In the other regions, a two-finger touch and drag will move a document. We reserved the one-finger touch and triple-finger touch for actions that map more naturally to the associated operations, as explained below.

In the Cutting Region, a one-finger touch will cut an image horizontally along a straight line extending from the finger touch position to both vertical edges of the image as in

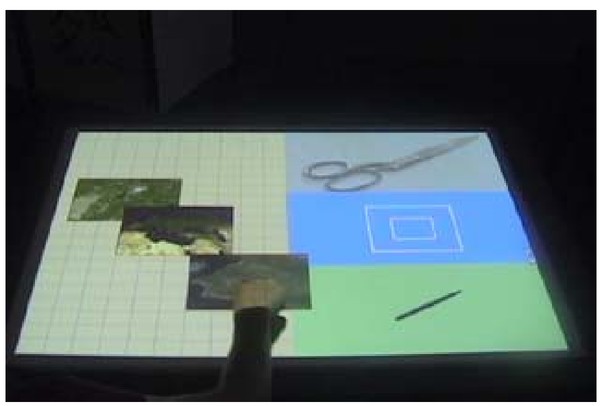

Figure 1: A touch sensitive tabletop surface with 4 Modal Spaces – cutting space (top right), resize space (center right), annotation space (bottom right), and layout space (left side).

Space	1 touch	2 touch	3 touch
Layout	Move	Move	Move
Cutting	Horizontal cut	Move	Bounding box cut
Resize	Grow	Move	Shrink
Annotation	Annotate	Move	Clear

Table 1. Mapping of three gestures to four Modal Spaces.

Figure 2. A triple-finger touch will cut out a rectangular portion of the image. The distance of the two fingers whose touch points are farthest apart defines the two diagonal corners of the rectangle to be cut out. Figure 3 shows a portion of the top image being cut out this way.

In the Resize Region, a one-finger touch will enlarge the image, while a triple-finger touch will shrink an image. In the Annotation Region, a one-finger touch and drag will make digital marks on the images, while a triple-finger touch will erase any marks made.

As there are only three different gestures, users can learn these quickly while having access to seven different operations: move, cut, cut-bounding-box, shrink, grow, annotate and erase. Users need only learn the mappings for each region, and they do not need to learn new gestures for every action.

DISCUSSION

An important consideration in interface design is determining an appropriate mapping from user actions to commands. A particular action might easily be used for a number of different commands. For example, single-finger interaction seems to lend itself to a number of possible actions (e.g., selecting, annotating, and moving). By using a modal interface we can overload a user's action, choosing which command to execute based on the location of the object and the action, thus simplifying the interaction.

Another important issue is direct-touch input and gesture design. Modal Spaces allow the user to leverage the power of gestures without having to memorize a large gesture set. This paper introduces the notion of Modal Spaces; we leave the design of a good gesture set to future work.

Large direct-touch displays have a different form factor than the desktop. Solutions that work well on the desktop, like the task bar, may not make sense on large direct-touch displays. For example, some users cannot physically reach the top of the screen in a vertical display, and there is no consistent "top" at all on an interactive tabletop.

Modal Spaces have an advantage over other methods in that they do not require the time or shift of focus necessary to explicitly select the mode. They are well-suited to use over multiple surfaces, as they provide a feasible way to perform multiple different actions without requiring users to search for controls at every surface.

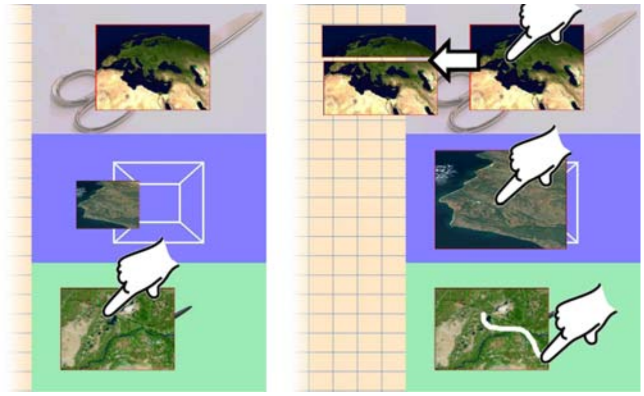

Figure 2: A before-and-after look at the effects of one-touch input interaction: cutting, growing and annotation.

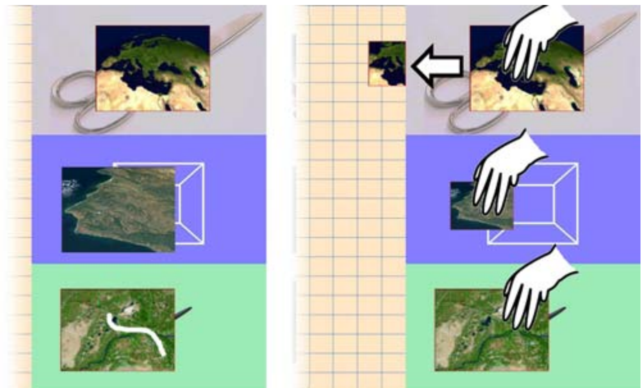

Figure 3: A before-and-after look at the effects of a triple-touch: bounding box cut, shrinking and clearing annotations.

Our implementation is document-centric as opposed to application-centric: documents carry their type with them so the user can arrange the documents in as many ways as he/she needs to. This is important for a large display because multiple people can use the display at the same time and accessing documents is more flexible than managing applications.

Modal Spaces is not a technique intended to replace the more complex operations enabled by menubar and toolbar functionalities. Rather, it is intended as an in-between technique, an enhancement and complement for laying out interactions that could benefit from spatial semantics.

Potential Applications

Modal Spaces afford a way of changing a document's view based on its location. This is particularly useful on a large screen, and provides comparable functionality to Magic Lenses [1], which have been used on traditional smaller screen displays. If the goal of a space was to identify and mark steep regions on a map, then placing a map in that modal region may display it as a set of topological lines rather than an image. If the task is looking for specific geological features, then another modal region could display the map as a pixelated image. This way the user can move the map among the spaces, keeping her focus on a particular feature and not losing track of what she is looking at while looking for a button or menu item to change views. She can easily compare how a map in topological form compares to one in visual form by dragging it into the appropriate modal region. If the maps need certain regions highlighted, landmarks indicated, and scale adjusted, modal spaces would allow a user to flexibly move maps between marking, highlighting, and adjusting regions. This allows the user to see and work with different maps in parallel while keeping her focus on the maps rather than tool selection.

The domain of simultaneous video coding of several video streams of the same event from different perspectives can also benefit from Modal Spaces. On some videos the user may want to mark participant utterances, on-screen events, or user action. The experimenter can classify the video streams by dragging them into the different modal regions, thus changing what his annotations will show. The marks will then be organized into different types without requiring him to change a tool as he interacts with different video streams. For example, in one modal region a touch could indicate audio utterance begin and end, and in another gaze direction.

Design Issues and Future Research

In this section, we discuss issues that we have observed in early casual usage of our image modification application, as well as some general issues that have become apparent during the design of our proof-of-concept system.

Matching of Task to Space and Gesture to Task

During particular tasks, a user requires specific functionality available, and it is necessary that the gesture-action mapping reflects this. For this solution to be usable, it is important to carefully match the workspace with the task. When designing a Modal Space interface, it is critical to create affordances in the space to ensure that the gesture set is visible and clear.

Appropriate Display Form Factor

Modal spaces can be of use in several situations. However, one limitation to this technique is that it requires adequate screen space; thus it works best on large displays. With a large workspace, such as an interactive tabletop, some natural location-based partitioning already occurs [3]. This leads us to believe that it may be feasible to use the Modal Spaces concept as a solution to some of the challenges inherent in designing interactions for large displays.

Scalability and Layout

The number of spaces feasible at one time is limited. It is application and usability dependent. One possible solution is to provide multiple digital workbenches, and to allow the user to switch between them to change the regions available on the table. As in the case of the physical workbench, the layout of the spaces is very important. In a system of this type, it can be valuable to allow on-the-fly adjustment of the regions to maximize useful screen space and user control. One extension we are currently pursuing is a "stretchable" and "panning" Modal Spaces canvas on large interactive tabletops. A digital canvas allows the user to "pull" a modal space that is visible but not reachable closer to his/her immediate work area. Conversely, we are also exploring compression of regions and the documents within them into smaller areas. Choice of screen layout depends on what is most task-appropriate and comfortable for the user.

Multiple Users

One advantage of large displays is that they allow multiple people to interact simultaneously. In Modal Spaces, there can be a conflict if two people want to use the same space at the same time. Also, if a modal space is convenient to one user, it may not be reachable by another. Designing Modal Spaces for multiple users across a large surface, or even across multiple display surfaces is an interesting research challenge.

For future work, we intend to evaluate how Modal Spaces compare to other ways to indicate mode and explore how to integrate Modal Spaces with other interaction techniques.

CONCLUSION

In this paper we have discussed the difficulties involved when applying GUI-based techniques to solve the mode problem on large, direct-touch displays. We present a new technique, Modal Spaces, to address this challenge. The Modal Spaces technique addresses two key issues for direct-touch modal interfaces: indicating and changing mode in a modal interface. It supports visual identification of mode and allows reuse of the same gesture set. Modal Spaces can enhance and complement the current mode interfaces for larger display form factors, and open new avenues of research.

ACKNOWLEDGMENTS

We would like to thank our anonymous reviewers for their helpful feedback.

REFERENCES

1. Bier, E. A., Stone, M., Pier, K., Buxton, W. & DeRose. T. Toolglass and magic lenses: the see-through interface. *Proc. SIGGRAPH*, ACM Press (1993), 73-80.
2. Dietz, P. and Leigh, D. DiamondTouch: A Multi-User Touch Technology. Proc. of UIST 2001, 219-226.
3. Scott, S.D., Carpendale, M.S.T, Inkpen, K.M., Territoriality in Collaborative Tabletop Workspaces. *Proc. CSCW 2004*, ACM Press (2004) 294-303.
4. Wu, M., & Balakrishnan, R., Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. ACM UIST 2003, 193-202.