The Continuous Interaction Space: Integrating Gestures Above a Surface with Direct Touch

Nicolai Marquardt[1], Ricardo Jota[2], Saul Greenberg[1], Joaquim A. Jorge[2]

[1]Department of Computer Science
University of Calgary
Calgary, Alberta, CANADA, T2N 1N4

[2]VIMMI Group / INESC-ID
IST / Technical University of Lisbon
1000-029 Lisbon, PORTUGAL

{nicolai.marquardt, saul.greenberg}@ucalgary.ca, rjc@vimmi.inesc-id.pt, jaj@inesc-id.pt

ABSTRACT

The advent of touch-sensitive and camera-based digital surfaces has spawned considerable development in two types of hand-based interaction techniques. In particular, people can interact: 1) *directly on the surface* via direct touch, or 2) *above the surface* via hand motions. While both types have value on their own, we believe much more potent interactions are achievable by unifying interaction techniques across this space. That is, the underlying system should treat this space as a continuum, where a person can naturally move from gestures over the surface to touches directly on it and back again. We illustrate by example, where we unify actions such as selecting, grabbing, moving, reaching, and lifting across this continuum of space.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General terms: Design**,** Human Factors

Keywords: Surfaces, touch, gestures, proximity

INTRODUCTION

The advent of highly interactive digital surfaces has motivated researchers to develop a rich set of accompanying interaction techniques. While there are now a broad variety of techniques, *hand gestures* dominate. So far, most prior research has focused on two gestural modes.

a. *On the surface* includes touch interactions directly on the reachable parts of the display, usually using fingers or hands [1, 3, 5, 10, 13]. Interactions typically include selecting, grabbing, throwing, rotating and moving.
b. *Above the surface* includes pointing and hand gesture recognition at near proximity to – but not on – the surface. These interactions typically select and access content not reachable by the user's direct touch [8, 12], or map gestures to particular actions [4, 6, 11].

While previous research has explored the characteristics of these two gestural types, they are usually treated as distinct modes. Our perspective differs. Rather than having 'on the surface' *or* 'above the surface' interaction, we see this interaction space as a continuum. That is, a person's gestural touch directly on the surface should naturally flow into gestural acts above the surface, and back again. In turn the digital artifacts on the surface, as well as the feedback offered by the surface, should respond in ways that matches this continuous action.

Marquardt, N., Jota, R., Greenberg, S., Jorge, J.A. (2009) The Continuous Interaction Space: Integrating Gestures Above a Surface with Direct Touch. Report 2009-925-04, Department of Computer Science, University of Calgary, Calgary, AB Canada T1W 1S2. Includes video figure.

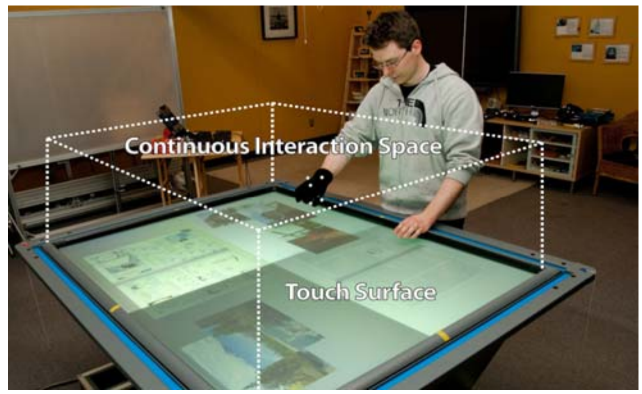

Figure 1. The continuous interaction space of a surface.

Our research goal is to understand how the space on and above the surface – including direct touch – can be unified for continuous interaction. This goal implies two sub-goals that we explain in the following sections of the paper.

1. Construct gestures that take full advantage of this unified space between touch on the surface and the space above it. In order to construct such gestures, we first define the concept of a *continuous interaction space* and describe why we think this is relevant for interaction.
2. Have digital objects respond in an appropriate way to the particular stage of gestural action in this continuum. This includes object behavior and feedback cues that support gesture continuity throughout the unified space.

To illustrate our ideas we present several example interactions with digital artifacts on a digital tabletop (see accompanying video); while specific to a horizontal display, the underlying concepts should generalize to other digital surfaces such as electronic whiteboards and large wall displays. Our prototype system is implemented with an interactive horizontal *SmartBoard*[1] with a touch sensitive surface, and a *Vicon*[2] marker tracking system. The Vicon is, of course, not practical for deployment out of the lab. However, we believe our examples of unified interaction techniques could be implemented on much more affordable technologies that are now emerging, e.g., shadow tracking [2], a switchable diffuser [4], or any other technology that can accurately detect movement above the surface.

[1] http://www.smarttech.com
[2] http://www.vicon.com

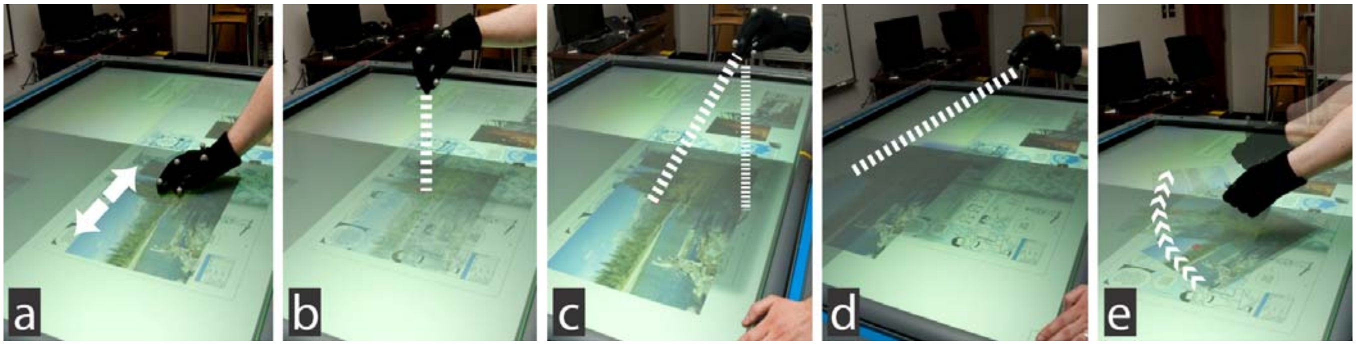

Figure 2. Continuous gestures for fluent interaction: (a) pick and drag, (b) revealing, (c) scaling, (d) moving, and (e) rotating.

THE CONTINUOUS INTERACTION SPACE

We define the *continuous interaction space* as being composed of the direct touch surface and the space above (bounded by ~1 meter), as illustrated and implemented in Figure 1. We argue that these are not two distinct spaces but instead a single interaction space. Specifically, a person can gesture fluently in this 3D area, where gestural acts flow from the space above, to touch, and vice versa. As well, we believe that gestures should not be limited to interactions immediately below one's hands. That is, gestural acts should *extend the space* to those areas of the surface that are physically out of reach. This view of a continuous interaction space suggests four interaction categories.

Distinct gestures either *work with direct touch or the space above*. That is, there is an explicit set of gestures used when touching the surface, and another separate set of gestures performed above the surface. This is currently how most prior research techniques exploit the continuous space (if at all). For instance, a touch-based two-finger pinching gesture directly on the surface is now common for scaling objects [1, 7], while moving a hand vertically above the surface is a method for changing interaction layers [6, 11].

Mirrored Gestures are gestural pairs that *redundantly encode identical functionality in either space*. That is, a person can invoke the same action via a gesture either directly on the surface or the space above it. Such gestures may be different or similar. Our first example uses two different gestures to produce an identical resize action: touch-based pinching (as mentioned above) vs. a two-handed gesture that brackets an object via L-shaped fingers and thumb, and then stretches or shrinks the area to scale the object. Our second example uses two similar gestures to produce an identical action; in this case the two-handed bracketing gesture can be on or above the surface.

The next two categories exploit the space as a continuum.

Extended Continuous Gestures. An extended gesture *extends its functionality* by letting the person operate the gesture (and the corresponding action) through the continuous space. That is, a particular action can begin in a specific place, but can also move through space – from touch to area above the surface – which affects that action's behavior. For example, consider an extended gesture for moving an object around a surface. A person starts by touching an object on the surface, and – while still touching – moves the object by dragging. The novelty is that this action can be continued by lifting the hand into the 3D space above the surface. Now, the hand extends to a pointing ray, where the person continues to move the object, even to positions out of reach by direct touch. What is important is that these are not two separate actions. Rather, they are done as a continuous flow.

Proximal Continuous Gestures. Our final type of gesture is somewhat similar to an extended gesture, except that it also exploits hand distance above the surface and hand orientation with respect to the surface as additional input dimensions, possibly triggering multiple actions. That is, the gesture exploits the *complete space of its added dimensions*. This allows complex operations to be interwoven into a single gesture. Figure 2 illustrates an example. As with most touch surfaces, the person can select and move an object via direct touch and drag (Figure 2a). While doing this, they can also move their fingers together into a 'pick' posture, and then lift their fingers off the surface into a 'pick-up' gesture. As in real life, this lets them pick up the object above its current surface plane (Figure 2b). The picked-up object becomes increasingly transparent to reveal the other objects that are now underneath it (transparency is a function of the vertical distance of the hand from the surface). At the same time, if the hand moves laterally while still pointing downwards, the object will move to follow under the hand. Next, if the pick gesture is pointed elsewhere other than straight down, the underlying object will move to the spot on the surface being pointed to (Figure 2c). Other effects occur: transparency decreases as the pointing angle moves away from the surface normal, and the object's scale factor changes as a function of the hand's vertical proximity to the surface (Figure 2d). If a person rotates their hand, the object rotates as well (Figure 2e). At any time, the person can release the object at its current location, size and orientation by opening up ones fingers. They can also return to direct touch.

We stress that the key idea of the above example is the flow of hand gestures across this space as a continuum, from touch to moving above the surface. Part of doing so leverages input dimensions such as distance above the surface (where 0=touch), and hand orientation with respect to the surface. By designing such gestures, people can interact with digital content in a very rich, fluid and complex way.

IMPLEMENTATION: INPUT TRACKING

Our system depends on touch information returned by the *SmartBoard* surface combined with the tracking coordinates of the hand markers from the *Vicon* cameras. Our raw information includes the finger position(s) on the surface (including liftoff/touchdown information), tracking of the hand model which returns the hand's yaw, pitch, and roll angle and its position in the 3D space, and the position of the surface in the 3D space. From these, we compute the pointing vector of the fingers to the surface (ray casting), and the normal vector of the hand perpendicular to the surface. We calculate the intersection of these vectors to the surface in screen pixel coordinates (so we know exactly what spot and objects on the surface these vectors are referring to). We also compute the hand's position, direction and vertical distance relative to the surface. We recognize postures such as pinching by measuring fingertip distance from one another. At a higher level, particular gestures (such as those in Figure 2) register for updates of these hand movements.

To show how this works, reconsider the proximal continuous gesture in Figure 2. This complex gesture set is generic, as it can manipulate the behavior of a broad variety of digital objects. Object selection occurs by matches direct touches on the surface with underlying objects, or by pointing with the index finger from above (using the pointing vector to determine which object is pointed to). Detecting the pinching posture, whether enacted on or above the surface, initiates object grabbing. During object movement in the 3D space, the length of the normal vector (which is the distance of the hand from the surface) is mapped to scaling of object content. At the same time, object rotation is mapped to the hand's roll value. Object transparency is a function of the angle between the normal vector of hand and the pointing vector of the index finger (thus revealing objects underneath the grabbed object without moving them). Importantly, these actions are performed on the selected object by a continuous movement of the hand from touching to the space above the surface.

INTERACTION WITH DIGITAL CONTENT

We now illustrate how our idea of a continuous interaction space and its four gesture categories integrates into the broader ecology of surface interaction. We show how *mirrored*, *extended*, and *proximal continuous* gestures can handle specialized digital content, such as videos and digital books, and how this content responds to these actions.

While the proximal continuous gestures of Figure 2 work on any generic digital object, this is not enough. In reality, most digital objects require other interaction techniques that are specific to their content, and we have to be careful that our generic interaction techniques leave enough room in the gesture space to allow for these custom techniques. Two examples illustrate how this can be done.

Our first specialized digital object is a *digital book*. Here, we use a *mirrored gesture* for page turning, where the person can freely choose to use either a gesture on or above

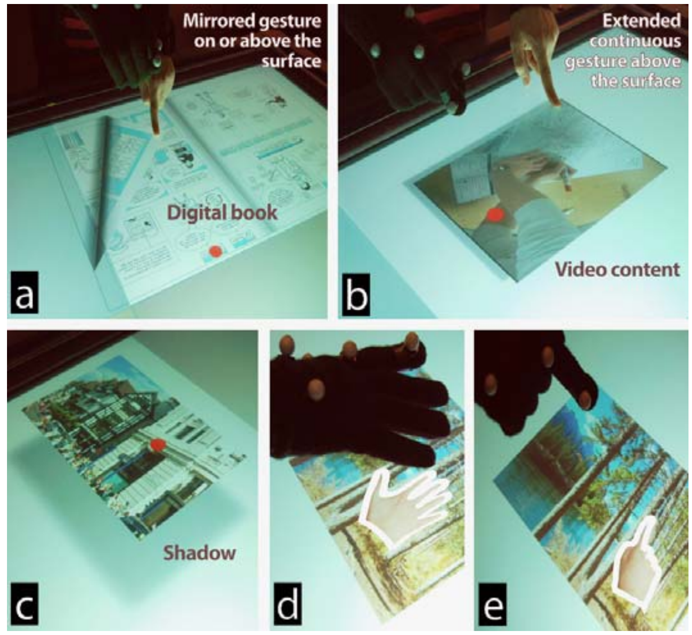

Figure 3. Interaction with digital content.

the surface to interact with the digital content. To flip a page a person either uses touch to flip a single page (Figure 3a), or does a gesture above the table that mimics page flipping. Depending on the speed, a user might turn multiple pages at once while performing this gesture in the space above the book. This works because it is the pinching posture that 'grabs' an object for the proximal continuous gesture of Figure 2, leaving the touch and page-flipping posture available for page selection and flipping.

Our next specialized digital object is a video object that can be searched and played. Here, we opted for an *extended gesture*. Using direct touch, a person can drag the finger to the left or right within the video object, which causes the video to advance forwards or backwards. Yet if the video object is small, this seeking becomes hard because small touch movements will translate to large jumps in the video timeline. We solve this problem by extending this gesture to the space above the surface (Figure 3b). Thus a person can initiate and begin the search by first touching the video object, and then refine the search by moving from touch to the space above the table; vertical distance from the surface is mapped onto the size of the jumps. Both lateral and vertical hand position controls the search (Figure 3b).

In summary, these examples illustrate how our proposed gestures combine a familiar interaction (usually touching the surface) with the additional 'space above' dimension for specialized control. Both still work alongside the proximal continuous gesture for generic object manipulation.

FEEDBACK CUES

To facilitate the interaction in this continuous space, digital objects should provide continuous feedback about their status. For instance, objects that are picked up from the table (with a grab gesture) and moved around by the user render a shadow onto the table surface, where the shadow size depends on the current distance of the hand to the surface (Figure 3c). This is a very natural mapping of the position in the 3D space to the displayed content on a 2D surface. Of course, all other object behaviors should reflect fine-grained actions corresponding to gestural movement: rotation, transparency, and so on. If done well, a person should be able to understand, self-correct, and fine-tune their gestures to control the object in a meaningful way.

The system should also provide feedback of how it is recognizing a person's gestures. We use an abstracted hand shape to give the person feedback of the current hand state as recognized within the unified interaction state. Examples are a pointing hand, pinching gesture, or the flat hand to reflect user actions (Figure 3d and 3e). These representations also help the interacting person to identify possible actions that can be performed with the digital content (e.g., selection, translation, rotation).

RELATED WORK

Touch-related research is thriving. While most initial work was on detecting contact points (e.g., one or two fingertips), the current interest has shifted to whole hand interactions. Wu and Cao et al. present a touch surface that understands such whole hand touches [1, 13]. Epps et al. studied hand shapes use in tabletop gestures [3]; their study suggests the need for both a touch screen and computer vision-based gesture tracking and recognition, for applications that require a wide range of commands. Others have explored how hand gestures can control a large scale display from a distance [5, 7, 12], and the influence of virtual embodiments to increase awareness when interacting with digital surfaces [9].

Distance has been investigated as an input metric in tabletops. Echtler et al. tracked hand shadows to support hovering actions on a tabletop [2], mimicking the mouse hover action. Izadi et al. applied a switchable diffuser to the tabletop, which captures hand gestures above the tabletop [4]. Parker et al. studied how point, in addition to touch, could improve interaction by enabling users to get out-of-reach objects [8]. Our own research expands beyond this two-state model by understanding pointing and touching as part of the continuous interaction space above surfaces.

Similar to us, some researchers are adding meaning to the space above the surface, for instance by dividing it into specific interaction layers. Lucero et al. [6] define gestures that allow vertical movement in-between layers to organize piles of pictures. Subramanian et al. present a multi-layer approach to tabletop [11], where they also divide the space above the tabletop as layers and include touch as the lowest layer. Our definition of the continuous interaction space builds upon these systems, as the discrete interaction layers of the related work can be understood as parts of the continuous space.

CONCLUSION

We proposed the concept of a continuous interaction space above a digital surface, where people can fluently move from touch interaction to gestures above the surface. Our categorization illustrates possible interaction techniques that benefit from extending into this continuous space above surfaces. *Mirrored gestures* allow the interacting person to choose a preferred gesture for a similar action. The *extended continuous gesture* allows one to continue the gesture above the surface, where the benefit is an additional dimension for better control. Finally, *proximal continuous gestures* allow complex operations to be interwoven into a single fluid gesture. Our examples illustrated gesture compositions that make use of this extended space. Of course, more advanced combinations and extensions of such gestures are possible. We do not claim that our examples are ideal ones, as there is much left to do. Overall, we believe that the understanding and designing gestures that exploit this continuous space above the digital surface is beneficial for creating intuitive interactions with the digital content.

ACKNOWLEDGMENTS

Thanks to Canada's NSERC, iCORE and SMART Technologies, Inc. who provided partial funding. Rob Diaz-Marino packaged the underlying Vicon infrastructure for use in this project. Ricardo Jota was supported by the Portuguese Foundation for Science and Technology (SFRH/BD/ 17574/2004).

REFERENCES

1. Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., and Hudson, S. ShapeTouch: Leveraging Contact Shape on In-teractive Surfaces. *IEEE Int'l Workshop on Horizontal Inter-active Human-Computer Systems*, (2008).
2. Echtler, F., Huber, M., and Klinker, G. Shadow tracking on multi-touch tables. *Proc. ACM AVI '08*, (2008), 388-391.
3. Epps, J., Lichman, S., and Wu, M. A study of hand shape use in tabletop gesture interaction. *Proc. ACM CHI '06 Extended abstracts*, (2006), 748-753.
4. Izadi, S., Hodges, S., Taylor, S., et al. Going beyond the display: a surface technology with an electronically switchable diffuser. *Proc. ACM UIST '08*, (2008), 269-278.
5. Lee, J., Lee, J., Kim, H., and Kim, J. Gesture-Based Interactions on Multiple Large Displays with a Tabletop Interface. *Universal Access in Human-Computer Interaction. Ambient Interaction*. 2007, 936-942.
6. Lucero, A., Aliakseyeu, D., and Martens, J. Augmenting Mood Boards: Flexible and Intuitive Interaction in the Context of the Design Studio. *IEEE Int'l Workshop on Horizontal Interactive Human-Computer Systems* (2007), 147-154..
7. Malik, S., Ranjan, A., and Balakrishnan, R. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. *ACM UIST'05*, (2005), 43-52.
8. Parker, J.K., Mandryk, R.L., and Inkpen, K.M. Integrating Point and Touch for Interaction with Digital Tabletop Displays. *IEEE Comput. Graph. Appl*. 26, 5 (2006), 28-35.
9. Pinelle, D., Nacenta, M., Gutwin, C. and Stach, T. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. *Proc. Graph. Interf.,* (2008), 1-8.
10. Ringel, M., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. *ACM CHI '04 Extended Abstracts* (2004), 1441-1444.
11. Subramanian, S., Aliakseyeu, D., and Lucero, A. Multilayer interaction for digital tables. *Proc. ACM UIST'06*, (2006), 269-272.
12. Vogel, D. and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. *Proc. ACM UIST'05*, (2005), 33-42.
13. Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *Proc. ACM UIST'03*, ACM (2003), 193-202.