

The Continuous Interaction Space: Interaction Techniques Unifying Touch and Gesture On and Above a Digital Surface

Nicolai Marquardt¹, Ricardo Jota², Saul Greenberg¹, Joaquim A. Jorge²

¹ Dept. Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4

² VIMMI Group / INESC-ID IST / Technical University of Lisbon, 1000-029, Portugal

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

Abstract. The rising popularity of digital table surfaces has spawned considerable interest in new interaction techniques. Most interactions fall into one of two modalities: 1) direct touch and multi-touch (by hand and by tangibles) directly on the surface, and 2) hand gestures above the surface. The limitation is that these two modalities ignore the rich interaction space between them. To move beyond this limitation, we first contribute a unification of these discrete interaction modalities called the *continuous interaction space*. The idea is that many interaction techniques can be developed that go beyond these two modalities, where they can leverage the space between them. That is, we believe that the underlying system should treat the space on and above the surface as a *continuum*, where a person can use touch, gestures, and tangibles anywhere in the space and naturally move between them. Our second contribution illustrates this, where we introduce a variety of interaction categories that exploit the space between these modalities. For example, with our Extended Continuous Gestures category, a person can start an interaction with a direct touch and drag, then naturally lift off the surface and continue their drag with a hand gesture over the surface. For each interaction category, we implement an example (or use prior work) that illustrates how that technique can be applied. In summary, our primary contribution is to broaden the design space of interaction techniques for digital surfaces, where we populate the continuous interaction space both with concepts and examples that emerge from considering this space as a continuum.

Keywords: Touch, gestures, surfaces, interactive tabletops, 3D interaction, tangibles, portable devices, continuous interaction space.

1 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 input (sometimes holding tangibles) dominates. Most prior research has focused on two modes of recognizing hand input.

- (a) *On the surface* includes touch interactions directly on the reachable parts of the display, usually using fingers and hands [7] [10] [19] [40], or tangible objects [2] [32]. Interactions typically include selecting, grabbing, throwing, rotating, and moving.
- (b) *Above the surface* includes free-hand gesture recognition that occurs in the space above the surface. These interactions typically point to, select and access content not reachable by the user's direct touch [23] [24] [33], or map gestures to particular actions [17] [20] [30].

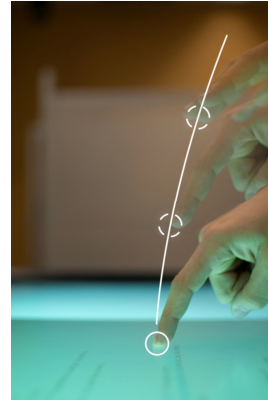


Figure 1. Continuous interaction from the space above the tabletop to direct touch.

The problem is that the vast majority of this prior research has explored hand interactions only within these two distinct modes. The limitation is that these two modalities ignore the rich interaction space between them. Our perspective differs. We propose and contribute a unification of these discrete interaction modalities called the *continuous interaction space*. The idea is that many interaction techniques can be developed to not only fit into a particular modality, but that they can leverage the space between them. That is, we believe that the underlying system should treat the space on and above the surface as a continuum, where a person can use touch, gestures, and tangibles anywhere in the space, and naturally move between them (illustrated in Figure 1).

In order to illustrate possible forms of interaction in the continuous interaction space, we introduce and contribute a variety of techniques that exploits the space between these modalities. We constructed these interactions in a way that takes full advantage of this unified space between touch on the surface and the space above it. For each interaction category, we implement an example or refer to prior work that illustrates how that technique can be applied. Thus, we also fit in earlier approaches of extending the interaction around digital surfaces; such as Hilliges' [16] and Wilson's [37] techniques for natural interaction with physically simulated digital objects, Baudisch's [2] stackable tangible blocks, or Subramanian's [30] interaction layers above a surface. While the techniques we introduce are specific to a horizontal tabletop display, the underlying concepts partially generalize to other digital surfaces such as electronic whiteboards and large wall displays.

In summary, our primary contribution is to broaden the design space of interaction techniques for digital surfaces, where we populate the continuous interaction space both with concept categories and examples that emerge from considering this space as a continuum. We don't claim that our categories are complete or our examples perfect. Rather, they show what is possible by considering the entire interaction space. The following section briefly introduces our understanding of the continuous interaction space. We then introduce our infrastructure that allowed us to develop and explore our interaction techniques. The main section of the paper introduces various categories of interaction techniques that illustrate how the unified interaction space can be applied in practice. We close with related work and a conclusion.

2 The Continuous Interaction Space

We define the *continuous interaction space* as being composed of the direct touch surface and the space above, as illustrated and implemented in Figure 2. We argue that these are not two distinct spaces but instead a single interaction space. Specifically, a person can interact 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, while the possible physical reach of a person's hands in this space is bounded to around ~1 meter above the surface, gestural acts can extend one's reach beyond these physical limits.

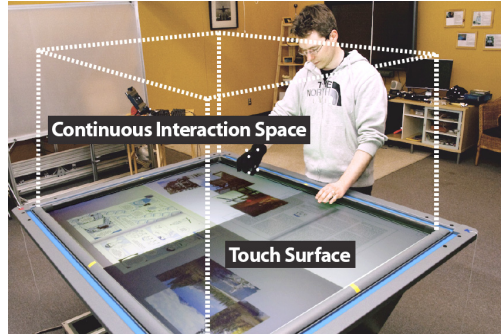


Figure 2. The continuous interaction space above the interactive tabletop surface.

3 Infrastructure for Rapidly Prototyping Interaction Techniques

In order to explore possible interaction techniques in the continuous interaction space, we built an interactive tabletop system infrastructure for sensing touch and gestures. Atop this, we developed an application test bed that contains a variety of objects that people can manipulate. In particular, a person can view collections of digital content (visible on the tabletop surface in Figure 2): text documents, photos, videos, as well as abstract entities such as files and folders. The person can interact with this digital content through a variety of techniques, such as moving digital content, creating stacks of items, turning pages in digital documents (Figure 3), picking up content to reveal documents underneath (Figure 4 left), navigating videos (Figure 4 right), or browsing stacks of documents (Figure 5). These are part of the interaction techniques that we explain in Section 4¹.

The underlying infrastructure is implemented atop an interactive horizontal touch-sensitive SmartBoard surface [www.smarttech.com]², and a Vicon motion tracking system [www.vicon.com]. This tracking system is composed of 8 high-speed infrared (IR) cameras that track reflective markers illuminated by IR spots attached to the

¹ All except two (4.7 and 4.8) of the interaction techniques we present have been implemented in our test bed application. We included these two additional techniques as they fit into our categorization of the continuous interaction space.

² In order to allow displaying multiple legible documents simultaneously, the projection surface of this interactive table prototype supports a high resolution of 2800x2160 pixels.

cameras. As long as at least three cameras of the system see each individual marker, the system can infer the accurate 3D position of this marker (to around 1mm). We attached markers to gloves to track people's hands on and above the tabletop surface (see right hand of the person in Figure 2). Our raw information includes the tracking of the hand model which returns the hand's and finger's yaw, pitch, and roll angle and its position in the 3D space (where a distance of 0 to the surface = touch), 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 (introduced shortly) register for updates of these hand movements.

We also track physical objects on and above the surface, where we attached similar IR reflective markers to tangible objects like wooden boxes (Figure 6), or portable tablet computers (Figure 5). Similar to the earlier described tracking of a person's hand, the system provides accurate position information of these tangibles. By comparing the distance of the tangible objects to the surface or a people's hands, we know whether the objects are placed on the surface or held in a person's hand.

Systems such as the Vicon motion tracking are currently too expensive and unwieldy for deployment out of a lab environment. However, we believe our examples of unified interaction techniques will be implementable in the near future on much more affordable emerging technologies, e.g., shadow tracking [9], switchable diffusers [16], depth cameras [37], or any other technology that can accurately detect position and movement above the surface. Currently, these other technologies do not yet provide the necessary accurate tracking of 3D positions of hands, fingers, and objects above a tabletop surface. Our use of marker-based motion tracking allows us to rapidly and accurately prototype and explore possible interaction techniques, where we can re-implement them when these other technologies are ready.

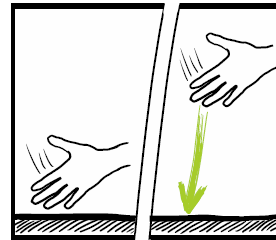
4 Interaction Categories

To illustrate the possibilities for novel interaction techniques in the suggested continuous interaction space, we now introduce a variety of categories of interaction techniques that make use of this extended space. We explain the concept behind each category. We also introduce one or more examples applying this technique: examples are mostly implemented in our own tabletop system, but a few refer to point systems found in the prior art. Our concepts begin with gestural interaction techniques of a person's hand with digital content. As we will see, some categories not only consider the position and movements of people's hands above the surface, but also people's use of tangible objects or digital devices (usually held in one's hands or placed onto the surface) for interacting with digital content displayed on the interactive surface.

Importantly, we stress that the key idea and contributions of the interaction categories explained below is the flow of hand gestures across the interaction 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.

4.1 Mirrored Gestures for Redundancy

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. Mirrored gestures are related to the concept of *equal opportunity* in interface design, as suggested by Cockburn and Bryant [8]. Mirrored gestures offer *multiple representations* for gesture input, adding to the flexibility of the tabletop interface. The redundancy of functionality lets a person select the interaction technique most convenient in a particular situation, and freely choose to use either a gesture on or above the surface to interact with the digital content.



The mirrored gestures may be *different* or *similar*. Our first example uses two different gestures to produce an identical resize action: touch-based pinching with fingers touching the surface *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. A third example considers the navigation through pages in a digital book. To flip a page a person either uses touch to flip a single page through dragging (Figure 3), or does a hand waving gesture above the table that mimics page flipping.

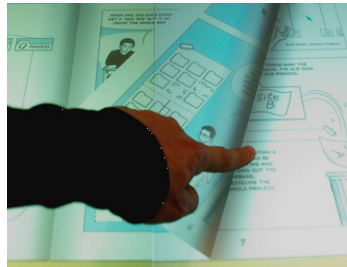
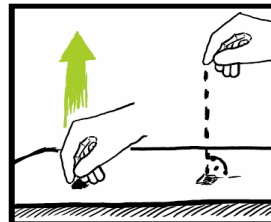


Figure 3. Turning pages in a digital document through direct touch.

4.2 Extended Continuous Gestures to Avoid Occlusion

With *extended continuous gestures*, a gesture that a person starts through direct touch on the interactive surface can *continue in the space above* the surface to avoid occlusion of the digital content visible on the tabletop display. That is, the person can lift the hand that performs the gesture off the surface, and continue the gesture in an invisible layer parallel to the touch surface.

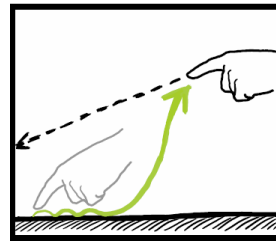


As an example, this technique can be applied to allow precise movement of small digital content displayed on the tabletop. A person can grab the object through a

'pick' gesture while touching the surface and move it along the tabletop plane through dragging. The novelty is that this action can be continued by lifting the hand into the 3D space above the surface. Now, this reveals the exact position of the digital object on the screen that was previously occluded by the fingers touching the surface. Once moving the hand above the surface, a normal vector of the plane through the pinching fingers defines the moving position of the object on the screen. In this technique, the actual distance of the person's hand to the surface has no impact on the movement action that is performed.

4.3 Extended Reach / Raycasting Gestures

The *extended reach* or *raycasting* technique allows a person to extend their interaction range – which is physiologically limited by the reach of their hands – to remote locations on the surface. The fingers of the hand or the arm of a person extend as a pointing ray when the hand is lifted above the surface, where the intersection of this pointing ray with the tabletop plane defines the point of action in the interface. This approach is related to techniques such as Grossman's raycasting for volumetric displays [12] or Parker's *TractorBeam* pen raycasting technique on tabletop interfaces [23]. It differs in the way this technique allows fluid transitions from dragging through direct touch to raycasting remote pointing when lifting the hand off the screen, and back to direct touch.



For example, consider an extended reach gesture for moving an object around a surface. Similar to the extended continuous gesture we introduced before, a person starts by grabbing an object with a 'pick' gesture on the surface, and – while still touching – moves the object by dragging. Lifting the hand off the touch surface *affects that action's behaviour*. 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.

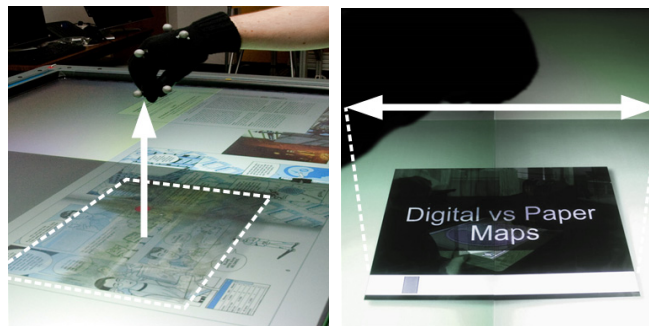
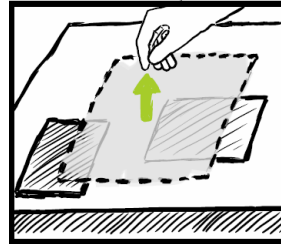


Figure 4. *Lifting up content to reveal documents underneath (left); navigating video with increased precision (right).*

4.4 Lifting Gestures to Reveal Objects

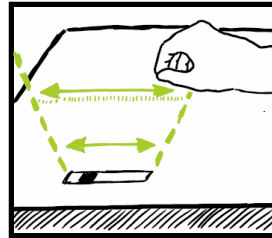
The *lifting gesture* allows one to virtually lift up digital content, primarily to reveal other content lying underneath. Our example has a person interacting with a large collection of digital documents. To reveal content that is currently occluded by overlapping documents, the person can move their 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.



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 (Figure 4 left).

4.5 Lifting to Adjust Scale Precision

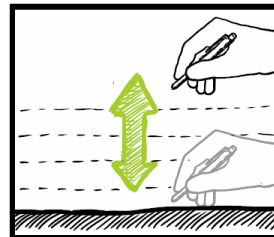
Our next interaction technique allows *lifting to adjust scale precision* to increase precision of lateral actions. Here, instead of just using the two dimensional touch surface as input to control a drag operation (e.g., moving a slider), a person can move their hand in the space above the surface to gain increased precision, where the lifting of a person’s hand in the Z-axis acts as an input parameter that readjusts the scale precision of the lateral movement.



To illustrate, consider a digital video object that can be searched and played. Using direct touch, a person can drag the finger to the left or right on a video navigation slider, which causes the video to advance forwards or backwards. Yet if the object is small (or the video large), navigating is hard because small touch movements translate to large jumps in the video timeline. With lifting to adjust scale precision, a person initiates and begins the search by first touching the video slider, and then gains precision by lifting above the table (Figure 4 right). Hand height continually rescales the slider, where the same lateral movement will result in a smaller move through the video. Thus both lateral and vertical hand position controls the search.

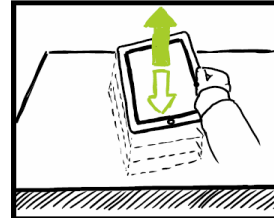
4.6 Interaction with Discrete Layers and Stacks of Digital Objects

Subramanian [30] previously suggested that the space above the tabletop surface can be divided into *discrete interaction layers*. In our adoption of this idea, the space above the surface is divided into multiple parallel planes, each corresponding to a layer. Each of these layers can then (for example) correspond to layers of visual content in the tabletop application, or even to different interpretations of gestures within a layer (e.g., an annotation layer, an editing layer, a movement layer, etc.). A person can navigate through these layers by moving in the Z-axis above the table. In our first example, a person uses a ‘pick’ gesture with their fingers (or use a



pen as in [30]) to select content in each of these layers, and even move content from one layer to another.

Our second example reconsiders layers as stacks. Here, a person browses the digital content ordered in a *stack of digital objects* (such as photos) by holding and moving a tablet computer in the space above the stack. The screen of the tablet computer then shows the particular photo corresponding to its particular Z position inside of the virtual stack of photos (Figure 5).



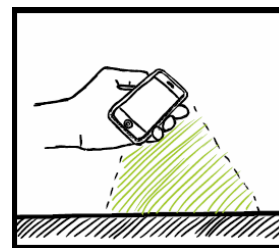
This technique has two practical limits. First, the maximum height of the virtual stack of documents is limited by the reach of a person's arms. If the person is holding a tablet (as in our 2nd example), this also limits the visibility of the tablet computer display. Second, the height of each individual layer is limited by the ability of people to select (and stay) inside of these layers. Subramanian's [30] evaluation found a layer thickness of 4 cm optimal for tasks where people have to stay in the layer. In our own implementation of browsing through a stack of documents with a tablet computer, we found a layer thickness of 1.0 cm to 1.5 cm as still practical for selecting content (and holding the device in this position).



Figure 5. Browsing stack of photos by holding and moving a tablet computer.

4.7 Magic Lenses and View Ports

Magic lenses (inspired by Bier et al. [5]) let a person move a plane above the tabletop to see individual, modified views onto the data that is visualized on the table [17] [27]. This moving plane can be, for instance, a sheet of paper that shows information of a projector (e.g., [17]), or a portable tablet-sized device that shows information. While moving these planes through the 3D space above the surface, the content displayed on the magic lens represents the current view corresponding to the 3D position of the lens. For instance, when visualizing 3D volumetric medial scans on the tabletop, the magic lens can visualize cut sections of that dataset [17].



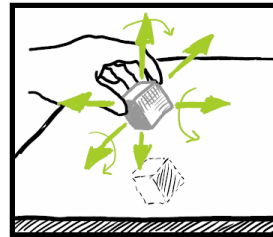
4.8 Stacking of Physical Objects

People can stack tangible objects on a tabletop, where the surface application reacts to the stacks built on the table. Baudisch et al. [2] proposed this approach with the *Lumino* project, where a Microsoft Surface tabletop recognizes stacks of small cuboid blocks. Bartindale et al. proposed a related approach detecting stacks of items on a surface, by determining the order of fiducial markers [1]. Overall, these approaches allow a person to create physical stacks and structures on top of the usually flat tabletop surface.



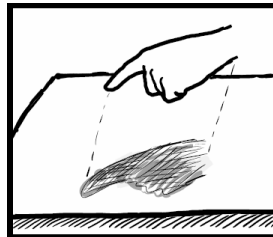
4.9 6-DOF Manipulation

The free movement of people's hands in the space above the surface can be leveraged to allow full *6 degree of freedom (DOF) interaction* with digital content. These dimensions are movement along and rotation around the three axes. In previous work, Hancock contributed 2- and 3-touch techniques that allow 6-DOF manipulations of objects directly on the surface (e.g., [13, 15]). Our technique differs: instead of using only multi-touch techniques on the surface, a person can 'pick up' any digital object, and modify the object's position and orientation by the way they move their hand above the surface. Our example displays a digital object (a 3D cube) that can be picked up; the cube's position and orientation are then directly mapped to the position of the person's hand (3D location: yaw, pitch, roll). Of course, the movement along this 6-DOF is again limited by the person's reach of the hand, and the rotation is limited by the movement around the wrist joint.



4.10 Feedback of Hand and Object Actions by Shadows

While it is relatively easy for a person to know what they are manipulating when they are directly touching it, this becomes harder to do when one's hand is lifted away from the surface as the physical connection is broken. To ease this, the system can provide *feedback* of how it is recognizing a person's gestures via *hand and arm shadows* (such as in [16] or [26]), or by visualizing abstracted hand shapes atop the objects it is manipulating. Examples of such shapes are a pointing hand, a pinching gesture, or the flat hand to reflect user actions atop the object(s) being manipulated. To avoid occlusions of user interface elements, transparency can be used to overlay feedback visualizations.

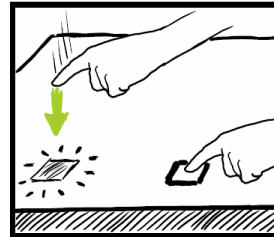


Digital objects can also provide additional continuous feedback about their status. For instance, objects that are picked up from the table (with a grab gesture) and moved around by the user can render shadows onto the table surface, where the shadow size depends on the current distance of the hand to the surface – or maybe

even ‘merges’ with the shadow of a person’s hand as proposed in [16]. 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 behaviours 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.

4.11 Feedback of Possible Actions by Hovering

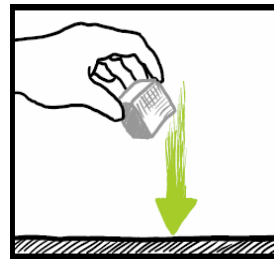
Related to the technique above is feedback about what *actions are possible* as a person moves and *hovers* above objects. For example, by sensing the position of people’s hands above the surface interface, widgets on the touch screen (such as buttons, sliders) can give people visual or audible feedback about possible actions (e.g., [9]). Similar to the way GUI buttons often change their color when the mouse cursor hovers over the button, the interactive elements of our tabletop application interface could change their visual appearance (e.g., glowing border, different background color) once a person’s hand approaches the widget. This technique would extend the common two-state interaction with touchscreens (i.e., touching the surface or not) to the three-state interaction (i.e., away, hovering, touching) familiar from GUI mouse interfaces or pen-based interaction [6], and possibly even to continuous states where the feedback is very specific to the possibilities allowed at a particular Z-coordinate.



As well, the hovering feedback does not necessarily have to be binary (*on* or *off*). Instead, widgets and other on-screen elements can adjust the feedback as a function of a person’s hand to the surface. For example, a button could begin showing a thin graphical border when the hand is around 20cm away, and continue increasing the thickness of this border the more the hand approaches the button. This kind of increasing amount of feedback might be in particular useful for widget actions with higher impact (e.g., deleting).

4.12 Picking and Dropping Gestures

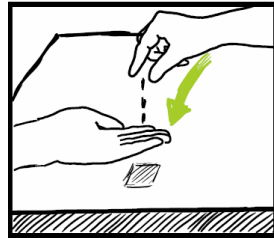
Another interaction technique in the continuous interaction space is *picking and dropping of digital objects*, where those objects react according to a physical simulation imitating real-world behaviour. This is also called *physics-based interaction* (e.g., [16]). That is, a person can pick up a virtual object (say a 3D cube) with a ‘pick’ gesture, move it to another location, and drop it again by releasing the gesture. The virtual objects displayed on the screen behave according to a simulated physical model and gravity; for example, they fall back to the ‘ground’, push other objects away, or stop when bouncing against ‘walls’. The Z distance can affect this behaviour, e.g., longer falls have larger consequences.



This approach of simulating the physics behaviour of digital objects has been the basis for various interactive tabletop systems (e.g., [3, 14, 16, 36, 37]). Some of the proposed interaction techniques are:

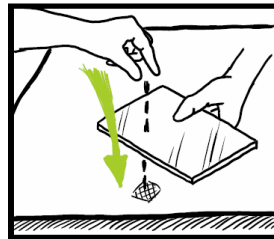
- Picking up objects through a ‘virtual joint’ between the hand and the object; placing objects in virtual containers (e.g., cup); piling and stacking virtual objects [16].
- Swiping objects off the side of the screen onto a person’s hand; moving the virtual content on a person’s hand to another surface [37].
- Picking and dropping virtual objects, augmented with 2D and 3D visualizations [3].
- Passing a virtual object from one person’s hand to another person’s hand (e.g., by dropping the object onto the other person’s hand) [37].

In these techniques, the distance of a person’s hand to the touch surface can have impact on the physical behaviour of the objects, such as when dropping an object and it falls back onto the surface. Depending on the modelling of simulation, the distance can affect the bouncing of an object (and thus the location where the object lands), or even the possible deformation of an object (e.g., when using rendered 3D shapes).



4.13 Picking and Dropping Objects Through Filters

As a novel variation of the above physics based interaction, we created another interaction category: *picking and dropping virtual objects through physical objects functioning as a filter*. In our example, a person picks a virtual object from the tabletop surface and lets the object fall through the filter (a thin acrylic panel). Once the virtual object ‘falls’ through the plane of the physical filter, the content of the digital object is modified according to the filter settings associated with this tangible filter object.



4.14 Pouring Gesture

When exploiting physics-based interaction, physical artefacts can function as containers for virtual objects, and afford gestures such as *pouring* to empty the container.

For example, in our prototype tabletop application a person can take a wooden box (Figure 6) and add virtual objects to this container by picking them up from the surface and dropping them from above into the container. The person can then move the box on the

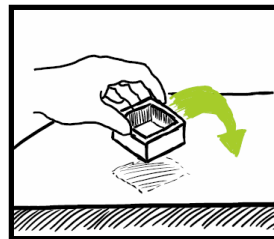


table or in the 3D space around the surface (the current virtual objects are displayed as shadows underneath the box). To release the content, the person turns over the box, and thus ‘pours’ the digital information out of the physical container (see Figure 6).

5 Related Work

Our work relates to previous research of interaction techniques out of the following three areas: multi-touch interaction with digital tabletop surfaces, 3D and physics based interaction, and tangible objects and digital devices on tabletops. In the following, we briefly review this work.

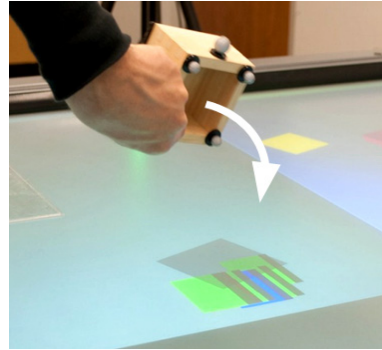


Figure 6. *Pouring* virtual documents out of a physical wooden box.

5.1 Multi-touch Tabletop Interaction

Touch-related research is thriving. While most initial work was on detecting contact points (e.g., one or two fingertips), current interest has shifted to whole hand interactions. Wu [40] and Cao et al. [7] present a touch surface that understands such whole hand touches. Epps et al. studied hand shapes use in tabletop gestures [10]; 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. Marquardt et al. introduced the fiduciary tagged glove on the Microsoft Surface that facilitates identifying the hand part that touched the surface [22]. Others have explored how hand gestures can control a large scale display from a distance [19, 21, 33], and the influence of virtual embodiments to increase awareness when interacting with digital surfaces [25].

Recently researchers began to use the space above the surface for interaction. Unlike the approach in this paper, most developed ‘point systems’ or techniques rather than consider the space holistically. Echtler tracked hand shadows to support hovering actions on a tabletop [9], mimicking the mouse hover action. Izadi applied a switchable diffuser to the tabletop, which captures hand gestures above the tabletop [17]. Parker’s Tractor beam [23] studied how pointing with a pen, in addition to touching the surface, could improve interaction by enabling users to get out-of-reach objects [24]. Benko’s muscle sensing technique for people’s hands also allows inferring information about their movement above the tabletop surface [4]. Other techniques include dividing the space above the surface into interaction layers. Lucero [20] defines gestures that allow vertical movement in-between layers to organize piles of pictures. Subramanian et al. present a multi-layer approach to tabletops [30], where people can interact with several layers of visual content by moving their pen in the 3D space above. Kattinakere [18] further investigated user performance of steering tasks along paths in such above-the-surface layers. We were inspired by (and thus integrated) concepts out of this body of related research into our categorization of techniques.

5.2 3D Gestural and Physics Based Interaction with Tabletops

A few researchers introduced physics based and 3D gestural interaction techniques to interactive tabletops. Wilson's *Touchlight* [39] and depth sensing above tabletops [36] opened up a field for rich 3D gestural interaction with digital content. Benko introduced cross dimensional gesture interaction techniques [3] that function as a bridge between 2D and 3D interaction. Grossman investigated interaction techniques with content on 3D volumetric displays [12]. Later, Hilliges introduced physics based interaction techniques enabling natural interaction with digital content [16]. This technique was later extended with depth-sensing capabilities [38], in order to provide an even richer set of possible gestures. In [37], this approach was taken further by including tracking of multiple interactive surfaces as well as the full 3D volume of the room they are located in. This permits interaction techniques spanning across multiple interactive surfaces.

Research of tabletop systems is partially inspired by other domains. These are, for instance, computer vision techniques for reliable tracking of people's hands and arms captured by cameras (surveyed in [41]), or techniques developed in virtual reality environment research for sensing 3D hand and object motions [29]: optical, marker-based, magnetic, and other tracking approaches.

Grossman's taxonomy of 3D on tabletops [11] categorizes properties of the visualization and interaction with such systems. In this context, we focus on tabletops with 3D volumetric direct input; but are not limited regarding display properties (2D, surface constrained, heads-up projection, or volumetric; see [11]). Our contribution is the emphasis on a continuous interaction space above surfaces, and the set of interaction techniques that in particular leverage the continuous input from people using touch, gestures, and tangibles on and above tabletop surfaces.

5.3 Tangible objects and digital devices on tabletops

One restriction when interacting with digital tabletops is the limited tangible expressiveness of a flat tabletop surface: even if the display visualizes 3D objects, a person cannot touch, grab, or feel these virtual objects. Consequently, researchers began to introduce tangible objects in tabletop interaction. In early research, Ullmer's *Metadesk* introduced physical interface elements in tabletop interaction [31]. Underkoffler's *URP* [32] combined tangible blocks representing buildings with a tabletop interface to enable intuitive interaction with a digital urban design simulation.

Starter's *Perceptive Workbench* [28] later combined the recognition of people's gestures above the tabletop surface with tangible computing. Later, Weiss introduced *Madgets* [35] as physical control widgets on a tabletop (such as buttons, sliders) that can automatically move through an emitted field of an array of electro magnets underneath the table. With Baudisch's *Lumino* [2], a tabletop system can also sense the geometrical form of physical cubes stacked on a tabletop.

Tangible objects or digital devices can also facilitate the exploration of visualized digital content. Inspired by [5], *magic lenses* allow a person to get a personalized view onto a data set, or cut sections of volumetric data visualized on the table [27]. In

another approach, Volda's *i-Loupe* allows a person use a mobile handheld device to view and interact with the content displayed on a large interactive surface [34].

While many techniques for tangible interaction on tabletop surfaces exist, our categorization focuses on those that in particular leverage the continuous distance of objects and their movements above the surface.

6 Summary

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. The idea is simple, yet the implications are profound. Most current interaction techniques are still anchored into falling in one of two modes: either based on direct touch *or* gestures above the surface. The concept of a continuous interaction space is to remove these arbitrary restrictions, where we want people to consider both modalities working together. Our categorization of techniques is suggestive, where we use them to illustrate an (incomplete) range of new interaction techniques that merge these modalities. Through example implementations and interwoven related work, we illustrated various gesture compositions that make use of this extended space.

Of course, more advanced combinations and extensions of such gestures are not only possible but highly probable. We do not claim that our categories or examples are ideal ones or exhaustive. 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. Our categories and examples illustrate some of the possibilities. As a starting point they suggest future exploration, iterative refinement, and eventual evaluation.

Acknowledgments

This research is partially funded by the iCORE/NSERC/SMART Chair in Interactive Technologies, Alberta Innovates Technology Futures, NSERC, and SMART Technologies Inc. Ricardo Jota was supported by the Portuguese Foundation for Science and Technology (SFRH/BD/ 17574/2004).

References

1. Bartindale, T. and Harrison, C. Stacks on the surface: resolving physical order using fiducial markers with structured transparency. *Proc. of ITS '09*, ACM (2009).
2. Baudisch, P., Becker, T., and Rudeck, F. Lumino: tangible blocks for tabletop computers based on glass fiber bundles. *Proc. of CHI '10*, ACM (2010), 1165–1174.
3. Benko, H., Ishak, E.W., and Feiner, S. Cross-dimensional gestural interaction techniques for hybrid immersive environments. *Proc. of IEEE VR '05*, (2005), 209-216.

4. Benko, H., Saponas, T.S., Morris, D., and Tan, D. Enhancing input on and above the interactive surface with muscle sensing. *Proc. of ITS '09*, ACM (2009), 93-100.
5. Bier, E.A., Stone, M.C., Pier, K., Buxton, W., and DeRose, T.D. Toolglass and magic lenses: the see-through interface. *Proc. of SIGGRAPH '93*, ACM (1993), 73-80.
6. Buxton, W. A three-state model of graphical input. *Proc. of INTERACT '90*, North-Holland Publishing Co. (1990), 449-456.
7. Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., and Hudson, S. ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. *Proc. of TABLETOP '08*, IEEE (2008).
8. Cockburn, A. and Bryant, A. Leogo: An Equal Opportunity User Interface for Programming. *Journal of Visual Languages & Computing* 8, 5-6 (1997), 601-619.
9. Ehtler, F., Huber, M., and Klinker, G. Shadow tracking on multi-touch tables. *Proc. of AVI '08*, ACM (2008), 388-391.
10. Epps, J., Lichman, S., and Wu, M. A study of hand shape use in tabletop gesture interaction. *CHI '06 extended abstracts*, ACM (2006), 748-753.
11. Grossman, T. and Wigdor, D. Going Deeper: a Taxonomy of 3D on the Tabletop. *Proc. of TABLETOP '07. IEEE* (2007), 137-144.
12. Grossman, T. and Balakrishnan, R. The design and evaluation of selection techniques for 3D volumetric displays. *Proc. of UIST '06*, ACM (2006), 3-12.
13. Hancock, M., Carpendale, S., and Cockburn, A. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. *Proc. of CHI '07*, ACM (2007), 1147-1156.
14. Hancock, M., ten Cate, T., Carpendale, S., and Isenberg, T. Supporting sandtray therapy on an interactive tabletop. *Proc. of CHI '10*, ACM (2010), 2133-2142.
15. Hancock, M., ten Cate, T., and Carpendale, S. Sticky tools: full 6DOF force-based interaction for multi-touch tables. *Proc. of ITS '09*, ACM (2009), 133-140.
16. Hilliges, O., Izadi, S., Wilson, A.D., Hodges, S., Garcia-Mendoza, A., and Butz, A. Interactions in the air: adding further depth to interactive tabletops. *Proc. of UIST '09*, ACM (2009), 139-148.
17. Izadi, S., Hodges, S., Taylor, S., et al. Going beyond the display: a surface technology with an electronically switchable diffuser. *Proc. of UIST '08*, ACM (2008), 269-278.
18. Kattinakere, R.S., Grossman, T., and Subramanian, S. Modeling steering within above-the-surface interaction layers. *Proc. of CHI '07*, ACM (2007), 317-326.
19. Lee, J., Lee, J., Kim, H., and Kim, J. Gesture-Based Interactions on Multiple Large Displays with a Tabletop Interface. In *Universal Access in Human-Computer Interaction. Ambient Interaction*. 2007, 936-942.
20. Lucero, A., Aliakseyeu, D., and Martens, J. Augmenting Mood Boards: Flexible and Intuitive Interaction in the Context of the Design Studio. *Proc. of TABLETOP '07. IEEE* (2007), 147-154.
21. Malik, S., Ranjan, A., and Balakrishnan, R. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. *Proc. of UIST '05*, ACM (2005), 43-52.
22. Marquardt, N., Kiemer, J., and Greenberg, S. What Caused That Touch? Expressive Interaction with a Surface through Fiduciary-Tagged Gloves. *Proc. of ITS '10*, ACM (2010), 139-142.
23. Parker, J.K., Mandryk, R.L., and Inkpen, K.M. TractorBeam: seamless integration of local and remote pointing for tabletop displays. *Proc. of Graphics Interface - GI '05*, Canadian Information Processing Society (2005), 33-40.
24. 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.
25. Pinelle, D., Nacenta, M., Gutwin, C., and Stach, T. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. *Proc. of Graphics Interface - GI '08*, Canadian Information Processing Society (2008), 1-8.

26. Shoemaker, G., Tsukitani, T., Kitamura, Y., and Booth, K.S. Body-centric interaction techniques for very large wall displays. *Proc. of NordiCHI '10*, ACM (2010), 463–472.
27. Spindler, M., Stellmach, S., and Dachsel, R. PaperLens: advanced magic lens interaction above the tabletop. *Proc. of ITS '09*, ACM (2009), 69–76.
28. Starner, T., Leibe, B., Minnen, D., Westeyn, T., Hurst, A., and Weeks, J. The perceptive workbench: Computer-vision-based gesture tracking, object tracking, and 3D reconstruction for augmented desks. *Machine Vision and Applications 14*, 1 (2003).
29. Sturman, D.J. and Zeltzer, D. A Survey of Glove-based Input. *IEEE Comput. Graph. Appl. 14*, 1 (1994), 30-39.
30. Subramanian, S., Aliakseyeu, D., and Lucero, A. Multi-layer interaction for digital tables. *Proc. UIST '06*, ACM (2006), 269-272.
31. Ullmer, B. and Ishii, H. The metaDESK: models and prototypes for tangible user interfaces. *Proc. of UIST '97*, ACM (1997), 223–232.
32. Underkoffler, J. and Ishii, H. Urp: a luminous-tangible workbench for urban planning and design. *Proc. of CHI '99*, ACM (1999), 386-393.
33. Vogel, D. and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. *Proc. of UIST '05*, ACM (2005), 33-42.
34. Volda, S., Tobiasz, M., Stromer, J., Isenberg, P., and Carpendale, S. Getting practical with interactive tabletop displays: designing for dense data, "fat fingers," diverse interactions, and face-to-face collaboration. *Proc. of ITS '09*, ACM (2009), 109–116.
35. Weiss, M., Schwarz, F., Jakubowski, S., and Borchers, J. Madgets: actuating widgets on interactive tabletops. *Proc. of UIST '10*, ACM (2010), 293–302.
36. Wilson, A. Depth-Sensing Video Cameras for 3D Tangible Tabletop Interaction. *Proc. of TABLETOP '07*. IEEE (2007), 201-204.
37. Wilson, A.D. and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. *Proc. of UIST '10*, ACM (2010), 273–282.
38. Wilson, A.D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., and Kirk, D. Bringing physics to the surface. *Proc. of UIST '08*, ACM (2008), 67-76.
39. Wilson, A.D. TouchLight: an imaging touch screen and display for gesture-based interaction. *Proc. of ICMI*, ACM (2004), 69-76.
40. Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *Proc. of UIST '03*, ACM (2003), 193-202.
41. Yilmaz, A., Javed, O., and Shah, M. Object tracking: A survey. *ACM Computing Surveys (CSUR) 38*, (2006).