

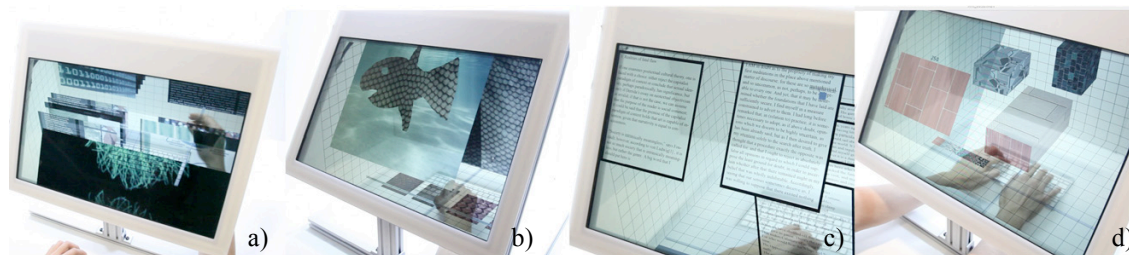
# SpaceTop: Integrating 2D and Spatial 3D Interactions in a See-through Desktop Environment

Jinha Lee<sup>1,2</sup>Alex Olwal<sup>1</sup>Hiroshi Ishii<sup>1</sup>Cati Boulanger<sup>2</sup>MIT Media Laboratory<sup>1</sup>Microsoft Applied Sciences Group<sup>2</sup>

75 Amherst St. Cambridge, MA, USA

1 Microsoft Way Redmond WA, USA

{jinhalee, olwal, ishii}@media.mit.edu, catib@microsoft.com



**Figure 1.** SpaceTop affords a) 3D direct spatial interaction, b) 2D direct touch, c) 2D indirect interaction, and d) typing. It aims to meld seams between these modalities by accommodating them in the same unified space and enabling fast switching between them.

## ABSTRACT

SpaceTop is a concept that fuses 2D and spatial 3D interactions in a single desktop workspace. It extends the traditional desktop interface with interaction technology and visualization techniques that enable seamless transitions between 2D and 3D manipulations. SpaceTop allows users to type, click, draw in 2D, and directly manipulate interface elements that float in the 3D space above the keyboard. It makes it possible to easily switch from one modality to another, or to simultaneously use two modalities with different hands. We introduce hardware and software configurations for co-locating these various interaction modalities in a unified workspace using depth cameras and a transparent display. We describe new interaction and visualization techniques that allow users to interact with 2D elements floating in 3D space and present the results from a preliminary user study that indicates the benefit of such hybrid workspaces.

## Author Keywords:

3D UI; Augmented Reality; Desktop Management

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI)

## General Terms

Human Factors; Design; Measurement.

## INTRODUCTION

Desktop computing today is primarily composed of 2D graphical user interfaces (GUI) based on a 2D screen with

input through a mouse or a touchscreen. While GUIs have many advantages, they can constrain the user due to the limited screen space and interaction bandwidth, and there exists situations where users can benefit from more expressive spatial interactions. For instance, switching between overlapping windows on a 2D screen adds more cognitive load than arranging a stack of physical papers in 3D space [7]. While there has been advances in sensing and display technologies, 3D spatial interfaces have not been widely employed in everyday computing. Despite advantages from spatial memory and increased expressiveness, potential issues related to precision and fatigue make 3D desktop computing challenging.

We present SpaceTop, an experimental prototype that brings 3D spatial interaction space to desktop computing environments. We address the previously mentioned challenges in three interdependent ways. First, SpaceTop accommodates both conventional and 3D spatial interactions in the same space. Second, we enable users to switch between 3D I/O and conventional 2D input, or even use them simultaneously with both hands. Finally, we present new interaction and visualization techniques to allow users to interact with 2D elements floating in 3D space. These techniques aim to address issues and confusion that arise from shifting between interactions of different styles and dimensions.

## RELATED WORK

Previous work has explored 2.5D and 3D representations to better support spatial memory in desktop environments [1, 7]. Augmented Reality systems exploit the cognitive benefits of co-locating 3D visualizations with direct input in a real environment, using optical combiners [8, 6, 5]. This makes it possible to enable unencumbered 3D input to directly interact with situated 3D graphics in mid-air [5, 9]. SpaceTop extends these concepts with an emphasis on streamlining the switching between input modalities in a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

unified I/O space, and the combination of such 3D spatial interaction with other conventional input modalities, to enable new interaction techniques.

Other related research explores the transitions between 2D and 3D I/O by combining multi-touch with 3D direct interaction [3], or through 2D manipulation of 3D stereoscopic images [4], with an emphasis on collaborative interaction with 3D data, such as CAD models. Our work focuses on how daily tasks, such as document editing or task management, can be better designed with 3D spatial interactions in existing desktop environments.

### SPACETOP IMPLEMENTATION

At first glance, SpaceTop looks similar to a conventional desktop computer, except the transparent screen with keyboard/mouse behind it. Users place their hands behind the screen to scroll on the bottom surface, or type on the keyboard. Through the transparent screen, users can view graphical interface elements appearing to float, not only on the screen plane, but also in the 3D space behind it or on the bottom surface. The users can lift their hands off the bottom surface to grab and move floating windows or virtual objects using “pinch” gestures.

We accommodate 3D direct interaction, 2D touch and typing, with an optically transparent LCD screen and two depth cameras (Figure 2) in a  $50 \times 25 \times 25 \text{ cm}^3$  volume.

#### Display: Prototype LCD with per-pixel transparency

We use a display prototype by Samsung, designed to show graphics without backlights in contact with its transparent LCD. The 22" transparent LCD displays  $1680 \times 1050$  pixels images at 60 Hz with 20% light transmission. It provides maximum transparency for white pixels, and full opaqueness for black pixels. We use the unique per-pixel transparency to control the opacity of graphical elements, allowing us to design UIs that do not suffer from the limitations of half-silver mirror setups, where pixels are always partially transparent. We ensure that all graphical elements include clearly visible opaque parts, and use additional lights for the physical space behind the screen, to improve the visibility of the user's hands and keyboard.

#### Head and hand tracking with depth cameras

One depth camera (Microsoft Kinect) faces the user and tracks the head to enable motion parallax. This allows the user to view graphics correctly registered on top of the 3D interaction space wherein the hands are placed. Another depth camera points down towards the interaction space and detects position and pinch-gestures of the user's hands [11].

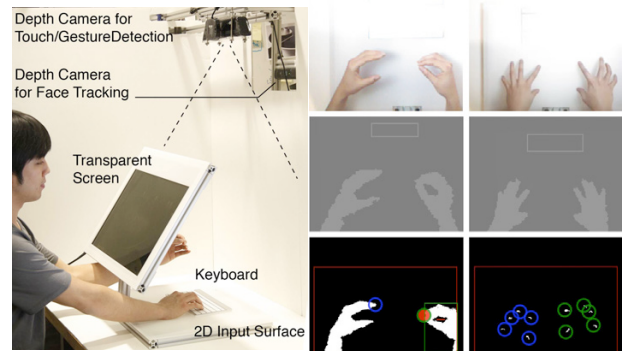


Figure 2. SpaceTop hardware configuration (left). Hand tracking and pinch detection (right).

The setup also detects if the user's hands are touching the 2D input plane based on the technique described in [10].

### INTERACTION AND VISUALIZATION

#### 2D in 3D: Stack Interaction

In SpaceTop, graphical UI elements are displayed on the screen or in the 3D space behind it. In our scenarios, details or 2D views of 3D elements are shown on the foreground plane (coinciding with the physical screen). While objects can take various forms in 3D space, we chose to focus on window interaction and 2D content placed in 3D space, such that the system can be used for existing desktop tasks. Another advantage of the window form factor in 3D is that it saves space when documents are stacked. It can, however, become challenging to select a particular window from the dense stack.

We designed various behaviors of stacks and windows to ease retrieval in stacks, as illustrated in Figures 3a–f. Users can *drag-and-drop* a window from one stack to another, to cluster it. As the user *hovers* his finger inside a stack, the layer closest to the user's finger gets enlarged and more opaque. When the user pinches on the stack twice, the dense stack *expands* to facilitate selection. The surface area below the stack is used for 2D gestures, such as scrolling. Users can, for example, *scroll* on the bottom surface of the stack to change the order of the documents in the stack.

We designed a *Grid and Cursor* system to simplify the organization of items in 3D. It provides windows and stacks with passive reference cues, which help guide the user's hands. The cursor is represented as two orthogonal lines parallel to the ground plane that intersect at the user's finger tips. These lines penetrate the grid box that represents the interaction volume, illustrated in Figure 3a.

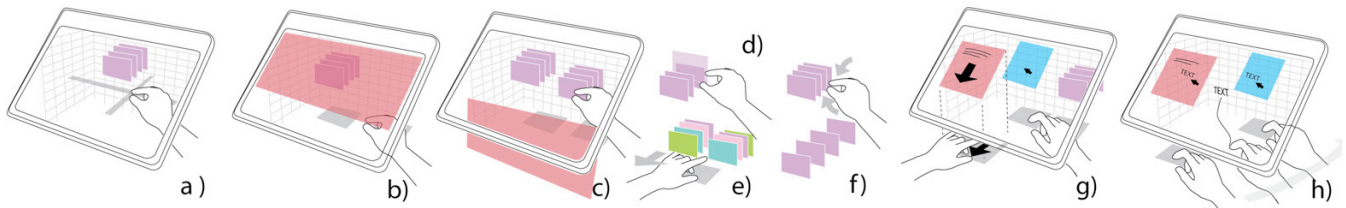


Figure 3. a) 3D grid and cursor. b), c) Sliding door. d) Highlighting. e) Scroll to change the order of windows. f) Expansion of a stack. g) Shadow Touchpad. h) Inter-shadow translation.



**Figure 4. Users can spatially manage windows in 3D space, and leverage spatial memory to retrieve tasks.**

#### *Modeless Interaction*

Our guiding principle for designing high-level interfaces and visualizations is to create a seamless and modeless workflow. Experiments have shown that when users shift from one interaction mode to another, they have to be visually guided with care, such that the user can mentally accommodate the new interaction model. Particularly, smooth transition between 2D and 3D views, and between indirect and direct interactions are challenging, since each of them is built on largely different mental models of I/O.

#### *Sliding Door: Entering the Virtual 3D space*

In the 2D interaction mode, the user can type or use a mouse or touchpad to interact with SpaceTop, as in any conventional 2D system. When the user lifts her hands, the foreground window slides up or fades out to reveal the 3D space behind the main window. When the hands touch the bottom surface again, the foreground window slides down again, allowing users to return to 2D-mapped input. The sliding door metaphor can help users smoothly shift focus from the “main” 2D document to “background” contents floating behind (See Figures 3b-c).

#### *Shadow Touchpad: One Touchpad per window*

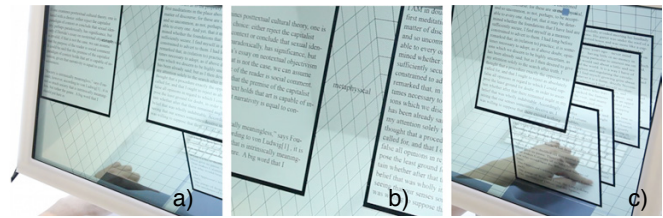
Touchpad interaction with 2D windows floating in 3D space introduces interesting challenges. Especially when working with more than one window, it is not straightforward how to move a cursor from one window to another. Indirect mapping between the touchpad and the window can conflict with the direct mapping that each window is forming with the 3D space. To address this issue, we propose a novel concept called Shadow touchpad, which emulates a 2D touchpad below each of the tilted 2D documents floating in 3D space. When a window is pulled up, a *shadow* is projected onto the bottom surface, whose area functions as a touchpad that allows the user to interact with that window. When multiple screens are displayed, each of them has its own shadow touchpad area.

#### *Inter-shadow translation of 2D element*

Users can move 2D objects (e.g., text and icons) from one window to another by dragging the object between the corresponding shadow areas. The object will be visualized as a floating 3D object during the transition between the two shadow touchpads, similarly to the balloon selection technique [28], as shown in Figure 5.

#### *Task Management Scenario*

Effective management of multiple tasks has been a central challenge in everyday desktop computing. In SpaceTop, the background tasks occupy a fixed position in the 3D space



**Figure 5. a) Selecting a piece of text and dragging it out of the source document. b) Text appears to be floating in open space while being moved from one document to another. c) The text appears again in the destination document.**

behind the main task, allowing users to rely on their spatial memory to retrieve them. This spatial persistence mitigates some of the cognitive load associated with conventional task management systems. Sliding door or stack interaction can be directly applied to categorize, remember, and retrieve tasks (Figure 4).

#### **Bimanual, Multi-fidelity Interaction**

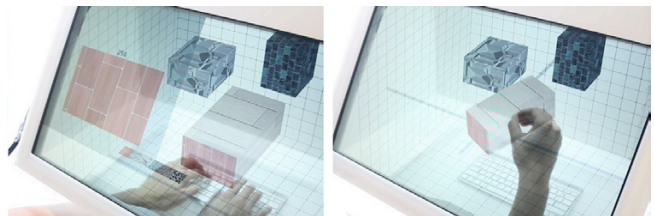
Interesting interactions arise when each hand is interacting in different styles and fidelities. The following applications demonstrate the potential of such bimanual, multi-fidelity interaction.

#### *Document Editing Scenario*

When composing a document, the user often needs to copy portions from other documents, such as, previous drafts or outside sources. SpaceTop allows the user to use the dominant hand to scroll through a main document, while simultaneously using the other hand to quickly flip through a pile of other documents, visualized in 3D space, to find a relevant piece of text. The user can then drag that piece of text into the main document through the more precise touchpad interaction. In this way, SpaceTop allows users to quickly switch back-and-forth between low-bandwidth, high-precision interactions (copying lines) and high-bandwidth, low-precision interactions (rifling through documents), or use them simultaneously.

#### *3D Modeling Scenario*

While 3D spatial interactions provide means for the user to materialize their design through spatial expression, much of the interaction in CAD require precise manipulation and is controlled in 2D. SpaceTop allows for natural transitions between these interaction modes. The user can start prototyping a model with free-form manipulation. Once fine control is required, the user can select a surface of the 3D model and pull up an editing console to the foreground



**Figure 6. Modeling Applications. The user can pull up a face to the foreground windows, edit it on the 2D screen, and manipulate in 3D space.**

of the screen. The user can then precisely modify dimensions by dragging a side, typing a number, or choose material properties by touching a 2D palette on the ground.

### PRELIMINARY USER EVALUATION

Ten participants (age 19–29, 2 female) were recruited from a university mailing list, none of whom had previous experience with 3D user interfaces. They were able to familiarize themselves with the system until they performed each action comfortably (3–6 min). The total experiment time for participants was between 70–80 min.

#### Switching between indirect 2D vs. direct 3D interaction

12 partially overlapping colored windows (red, green, blue, or yellow), containing a shape (triangle, square or star), were shown. Participants were given tasks, such as “grab the yellow square and point to its corners”, or “trace the outline of the blue triangle”. They performed four different, randomized tasks for three spatial window configurations, for a total of 12 trials for each of two blocks. The *SpaceTop block* used spatial window placement with head-tracking and participants used a combination of gesture, mouse and keyboard interaction, for constant switching between typing, 2D selection and 3D interaction. In the *baseline block*, windows were shown in the display’s 2D plane and only mouse and keyboard interaction was available.

Questionnaire responses (5-point Likert scale) indicate that the SpaceTop interactions were easy-to-learn (3.9). Participants did however find it slower (3.2 vs 4.2) and less accurate (3.2 vs 4.6) than the baseline. Users’ comments include “after I repeated this task three times (with the same arrangement), my arm starts moving towards the target even before I see it”, “switching to another window is as simple as grabbing another book on my (physical) desk”. Another user commented that the physical setup constrains his arm’s movement which makes him exhausted easier.

#### Text editing: Search and copy/paste

Participants skimmed the contents of six different document pages placed in the 3D environment. They were then asked to find a specific word and pick-and-drop it into the document on the foreground screen (see Figure 5).

Six participants commented that it felt compelling to be able to quickly rifle through a pile of documents with one hand while another hand is interacting with the main active task. One user commented: “it feels like I have a desktop computer and a physical book next to it”. “This feels like a natural role division of right/left hand in the physical world”. Three users reported that they had a hard time switching their mental models from 2D indirect mapping (touchpad) to 3D direct mapping (spatial interaction), which occurs when the user tries to drag a word out of a shadow.

### DISCUSSION

Users’ comments suggest that fast switching and bi-manual interaction provide compelling experiences, and that they can benefit from spatial memory (task 1). We also gained

some insights for future improvements. A few users also commented that they might perform better with a stereoscopic display, in addition to the aid of the grid and cursor. Although previous work indicates that stereoscopy has limited benefit over monoscopic display with motion parallax [5], we plan to also explore a stereoscopic version of SpaceTop. We think that the visual representation could be better designed to provide users with clearer guidance. While the current configuration allows us to rapidly prototype and explore interactions, we plan to improve ergonomics and general usability with careful design of the physical setup.

### CONCLUSIONS AND FUTURE WORK

SpaceTop is a concept that accommodates 3D and conventional 2D (indirect/direct) interactions in a single workspace. We designed interaction and visualization techniques for melding the seams between different interaction modalities and integrating them into modeless workflows. Our application scenarios showcase the power of such integrated workflows with fast switching between interactions of multiple fidelities and bimanual interactions. We believe that SpaceTop is the beginning of an exploration of a larger field of spatial desktop computing interactions and that our design principles can be applied to a variety of current and future technologies. We hope that this exploration offers guidelines for future interaction designers, allowing better insight into the evolution of the everyday desktop experience.

### ACKNOWLEDGMENTS

We would like to thank Akimitsu Hogge and John Weiss for assistance in implementation, and Steven Bathiche, Vivek Pradeev, Otmar Hilliges, Andy Wilson, Hrvoje Benko, Pranav Mistry, and Microsoft Applied Sciences for valuable discussions.

### REFERENCES

1. Agarawala, A., and Balakrishnan, R., 2006. Keepin’ it real: pushing the desktop metaphor with physics, piles and the pen. CHI ’06, 1283–1292.
2. Benko, H., and Feiner, S. 2006. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selections. 3D UI ’06, 79–86.
3. Benko, H., Ishak, E., and Feiner, S. 2005. Cross-dimensional gestural interaction techniques for hybrid immersive environments. VR ’05, 209–216.
4. Hachet, M., Bossavit, B., Cohé, A., and Rivière, J., 2011. Toucheo: multitouch and stereo combined in a seamless workspace. UIST ’11, 587–592.
5. Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A. 2012. HoloDesk: direct 3d interactions with a situated see-through display. CHI ’12, 1283–1292.
6. Olwal, A., Lindfors, C., Gustafsson, J., Kjellberg, T., and Mattsson, L. 2005. ASTOR: An autostereoscopic optical see-through augmented reality system. ISMAR ’05, 24–27.
7. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., and Dantzych, M. 1998. Data mountain: using spatial memory for document management. UIST ’98, 153–162.
8. Schmandt, C. 1983. Spatial input/display correspondence in a stereoscopic computer graphic work station. SIGGRAPH ’83, 253–261.
9. Treskunov, A., Kim, S. W., Marti, S. 2011. Range Camera for Simple behind Display Interaction. IAPR MVA ’11, 160–163.
10. Wilson, A. Using a depth camera as a touch sensor. ITS ’10, 69–72.
11. Wilson, A. 2006. Robust computer vision-based detection of pinching for one and two-handed gesture input. UIST ’06, 255–258.