Evaluating Gesture-based Desktop Projection Models in a 3D Environment

Gavin Ji Guo
gavinguo@cs.toronto.edu

Michael Margel
mmargel@cs.toronto.edu

Jessica Perrie
perrie@cs.toronto.edu

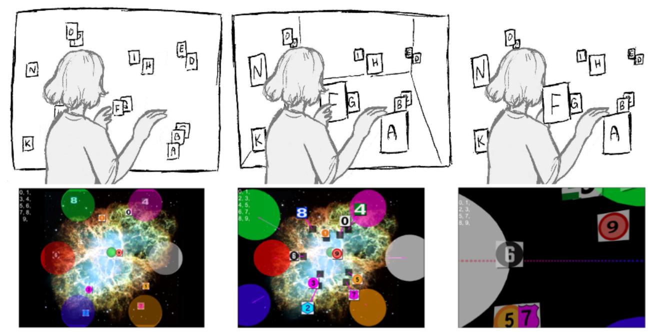
ABSTRACT
In this experiment, we attempted to leverage emerging augmented reality technologies in order to redefine the computer desktop paradigm. We introduce a new term, Desktop Projection Model (DPM), that is used to describe the way in which the desktop is projected into 3D space. By leveraging natural gestures and spatial memory, we hypothesized that a natural user interface, the immersive DPM (iDPM), would have better user performance and user experience than a more traditional interface, the planar DPM (pDPM). A within-participants study was conducted with 6 participants in order to determine whether spatial memory will allow users to complete tasks more quickly and more efficiently, and to determine which DPM yielded the best user experience. Data was gathered on the total time taken to complete each task, as well as the total distance that the participants needed to move their hands in order to complete each task. Our findings show that the iDPM required significantly more time and distance to complete each task, and that most participants preferred the pDPM – although most preferred the iDPM before tests. We believe that these findings were skewed by the technical limitations of the prototype, and that future study is needed to properly determine the viability of this interface.

Author Keywords
Gestural interaction; augmented reality; wearable computing; natural user interface; spatial memory; 3D UI; desktop management.

INTRODUCTION

Motivation
While computer technology has advanced to allow different multimodal input types and 3D graphics, desktop interfaces have generally remained the same – providing users with a 2D plane where items' and programs' icons can be arranged or organized as determined by the user using a traditional mouse, touchpad or touch-screen. We developed different 3D desktop projection models that enhance typical desktop interaction. By comparing these different models, we evaluate each's effectiveness to minimize the amount of time and amount of physical work needed in order for the user to manually cluster and relocate desktop objects.

Description
In this study, we define a Desktop Projection Model (DPM) as a way of projecting a desktop interface into 3D space, using an virtual reality system. The goal of this study is to explore how effectively users can manipulate different DPMs using gesture input, and to determine which DPM is the most efficient and the most preferred by users. User manipulation refers to organization and retrieval of desktop objects (e.g., document or program icons). We consider three different DPMs; each of them are sketched in Figure 1 and described below:



(1) Planar (2) Volumetric (3) Immersive

Figure 1. The three desktop projection models evaluated in this study, and their respective screenshots – interfaces visible to the user. Users interact with desktop objects using gesture input.

1. *Planar (pDPM)*: The pDPM consists of a vertical bounded 2D plane projected in 3D space. This DPM is to simulate a standard traditional 2D computer desktop floating in space.
2. *Volumetric (vDPM)*: The vDPM is to simulate a cube-like environment, which is the extension of a 2D desktop into 3D space. Specifically, each desktop object has an additional depth coordinate, corresponding to its distance from the front of the volume.
3. *Immersive (iDPM)*: The iDPM uses the space/room metaphor, consisting of the bounded outer shell of a sphere surrounding the user, rather than simply a space in front of the user. In the iDPM, the desktop objects occupy the physical space around the user.

Contributions
Unlike other studies [2, 4, 7, 14], we compare three different desktop projection models (DPM) based on perceived depth within a 3D interaction environment. We developed a 3D desktop interface visualization using the fundamental design principles of related literature [3, 17] and based on real-world metaphors for user interaction.

Our evaluation presents the results of three different DPM in terms of the amount of time and motion needed to complete the tasks, and based on user preference tested in a preliminary study. We analyse the results to identify which DPM yields the best performance, and discuss why this happens. Our findings are applicable to other areas for 3D interfaces: for example, navigating information visualizations of clustered data, or traversal of dynamic environments in game design.

Our paper first reviews the current state-of-the-art in 3D desktop interaction, and discusses the influences of natural user interfaces and spatial memory on our interfaces. We then outline our study to test each DPM, and discuss the results and their implications.

LITERATURE REVIEW

In developing and evaluating our 3D desktop interface, we rely on previous research about 3D desktop interaction and design, natural user interface (NUI), and support for spatial memory in a 3D user interfaces (UIs).

3D Desktop Interaction

Our design of the 3D desktop interface and interaction elements builds on previous work that developed BumpTop (3D desktop using a strong desktop metaphor) and SpaceTop (multidimensional interface with gesture input).

Agarawala et al. developed BumpTop – a 3D desktop interface that was designed to organize desktop objects using a pile metaphor (instead of filing) [1]. BumpTop is a 3D desktop interface that assigns desktop items with virtual physical characteristics such as weight, inertia, and friction. It aims at creating a virtual environment to make the user interaction with desktop objects more realistic – resembling real world desktops – much like how users would manipulate objects on a desk. BumpTop strictly follows the desktop metaphor; virtual boundaries are defined within a single viewpoint to simulate the real world desktop environment. To support and enhance the spatial memory effect, we extend this research by defining a larger boundary in the 3rd dimension (point inside/outside the display) while allowing users to change different viewpoints to navigate through the 3D space.

Filing and piling are two common strategies used to manipulate and facilitate desktop objects. They were studied and compared extensively in previous work [9, 10, 18]. However, most of these studies were conducted in the commercialized 2D desktop, which limited the possibilities in object organization.

Lee et al. introduced SpaceTop, a novel concept that combines 2D and spatial 3D interactions in a single desktop workspace [8]. Its design illustrates the possibility of switching from one modality to another, and of simultaneously using two modalities with different hands. SpaceTop allows users to interact with 2D desktop objects floating in a 3D space directly with hands (gestures) using depth cameras and a transparent display. Our work focuses on creating a 3D desktop space within a typical computer monitor, allowing users to store more objects, and improving usability (in terms of retrieval time, error rate, etc.) through spatial memorization.

Natural User Interface (NUI)

Natural User Interface (NUI) was first coined by Steve Mann as an interface approach to allow users to interact with the real world through natural actions (e.g., gestures, speech) as an alternative to command-line interface (CLI), graphical user interface (GUI) and tangible user interface (TUI) [11]. Wigdor refined this term by asserting that "a NUI is not a natural *user interface*, but rather an interface that makes your user act and feel like a natural" [19]. Thus, through interactions based on natural user actions, the NUI interface aims to provide a seamless transition from novice to expert user in terms of user experience.

Mistry et al. introduced WUW, a wearable gestural interface, which uses a small projector and a camera mounted on the user's hat to visually augment the physical world with interactive elements [13]. This design was developed further to become SixthSense, a computer-vision based wearable gestural interface that allows users to interact with the augmented physical world through natural hand gestures [12]. Both WUW and SixthSense recognize user gestures by detecting color markers attached to the user's fingers through a camera feed using computer vision. Additionally, users see the augmented physical world with digital information through a tiny projector mounted on their hat. To allow users to interact with the physical world using natural actions, we extend this research by replacing the gesture detection and recognition system with a depth camera (i.e., Leap Motion) mounted on a glass-like wearable display, allowing users to see the digital information directly as if it is part of the physical space as opposed to projecting the information on a physical surface (e.g., WUW, SixthSense).

Spatial Memory

Spatial memory is the user's ability to remember where an item is located [15]. Researchers have attempted to leverage this ability by provided various cues, dimensions, transitions and designs to produce faster results for spatial tasks [2, 3, 4, 7, 14, 15, 16, 17]. Our study aims to build on this research by comparing different desktop projection models (DPMs) within a 3D environment.

Robertson et al. presented Data Mountain, a visualization approach that allowed users to retrieve Internet Favourites using web page screenshots placed on a front-facing 3D "mountain", a helpful landmark to help users' navigate the virtual reality [15]. Data Mountain employed various cues make the UI appear 3D: continuous perspective sizing, occlusion, and appropriate sound effects. Data Mountain yielded faster retrieval task times than a typical text-based management system, which its use of spatial memory allowed for lower storage times, retrieval durations, and failure rates [15]. We extend this spatial metaphor by allowing multiple viewpoints and comparing different DPMs while Data Mountain and Task Gallery only used one viewpoint with mouse cursor input. We also incorporate landmarks (e.g., coloured spheres) to help users re-locate desktop objects.

Previous research has tested different inputs to interact with a spatially-based interface. Jetter et al. found that users demonstrated a stronger spatial memory when arm and hand movements (from multitouch input) were involved in a panning interface navigation (but not with panning and zooming) [7]. Rädle et al. found similar improvements to navigation based on spatial memory when egocentric body moments were used instead of traditional "drag-to-pan" and "pinch-to-zoom" touch gestures [14]. We extend these findings by employing arm and hand gestures to interact with the system.

In analysis between the usefulness of 2D versus 3D navigation, Andy Cockburn et al. found no significant difference in performance between the 3D and 2D interfaces [2, 4]. He also found that a third dimension for interaction actually hindered participants' performance [5]. While we compare dimensionality by testing different forms DPMs, we aim to provide a more immersive 3D environment with gesture-based input unlike Cockburn et al.'s studies, which only used monocular static computer displays with the more traditional mouse input [4].

Finally, spatial memory has been found to be better formed if the interface includes an element of difficulty while keeping use interesting and fun [3]. We aim to include both of these interaction designs in developing the 3D interface and test the interfaces in terms of time, interaction path, and user preference.

RESEARCH QUESTION AND HYPOTHESIS

Our study answers the following research questions:

1. Does an immersive space/room metaphor 3D desktop projection model (the immersive DPM) result in optimized user performance (less time and shortest virtual path taken) than other 3D projection models?

2. Does an immersive space/room metaphor 3D desktop projection model (the immersive DPM) result in higher user preference than other 3D projection models and traditional 2D desktop interface?

Our research questions are answered by applying our hypothesis on the experiment results (dependent variables) to compare the three different DPMs (independent variable). In accordance with related research that compared input methods [7], we will determine which DPM exhibits the best spatial memory performance (lowest task completion time), the best navigation performance (shortest physical path), and user preference (according to participants). Each of our hypotheses is outlined below:

H1. Spatial Memory Performance.

1. *The iDPM will have the best spatial memory performance.* More physical body-centric actions – potentially supported in this interface – have been shown to improve the time performance of navigation-based spatial tasks [14].

2. *The pDPM will have the second best spatial memory performance.* Although done with a 2D physical interface of displayed pages and selection using a laser pointer, Cockburn et al. found this interface yielded faster retrieval times than both physical and virtual (with mouse input) 3D-based interfaces [5].

H2. Navigation Performance.

1. *The pDPM will have the best navigation performance.* Its vertical 2D plane will have the smallest size compare to the other two DPMs.

2. *The iDPM will have the next best performance.* Applying the findings of [14], we believe that combined body and hand motion in efficient movement paths will yield better results than using gestures alone.

H3. User Preference.

1. *The iDPM will have the best user preference and will be the most natural/intuitive user interface.* The i DPM, compared to the other interfaces, functions in a way most supported by the Natural User Interface design, and hence, we believe that users will learn it quickly, and be impressed by its clear real-life metaphor.

2. *The vDPM will have the second best user preference.* Although a 3D interface yielded slower results, Cockburn et al. observed that users found it "natural" and rated it as being more effective than a 2D interface [4].

PILOT STUDY

In order to better understand the feasibility and effectiveness of the proposed DPMs, we first ran an initial within-participant pilot study. This allowed us to refine the methodology that would be used in the study.

Apparatus

To run the study, a number of materials were used as shown in Figure 2; to process and display the interfaces, we used a MacBook Pro with 4 GB of RAM, Intel i7 2GHz processor, and Mac OS. The system was connected to an Epson Moverio BT-100 Wearable Display (regulated using an Android-based controller) with 960x540 resolution, 60 Hz refresh rate, and a perceived image size of 80" at 5 meters. A prototype Leap Motion controller was used to track hand motions. A PhidgetSpatial 3/3/3 Basic accelerometer was attached to the wearable display during the experiment. In addition to this equipment, participants filled out an online questionnaire powered by Google, and were recorded using Lenovo YouCam video recording.

Participants

In the pilot study, we had one female participant, 24-years-old. She was recruited on a voluntary basis and expected the study to involve a 3D interface and wearable display. We required that all participants we use have average control of their arms and hands with either 20/20 vision or vision corrected to 20/20, and not to be color-blind.

Experimental Design

The independent variable in this study is the type of DPM being used, which has three levels: pDPM, vDPM, and iDPM. For each DPM, users were asked to complete two tasks that represented basic desktop management activities: replication (organizing the desktop) and retrieval (finding a set of desktop objects). Because the nature of the tasks are different, the results of each task cannot be compared against each other.

Apparatus Experiment set-up

Figure 2. Apparatus used in the experiment: a: Epson Moverio BT-100 Wearable Display, b: Leap Motion Controller, c: PhidgetSpatial 3/3/3 Basic accelerometer, d: Android-based Controller for Epson Moverio BT-100 Wearable Display, e and f: HDMI to AV converter. To use each DPM, participants sat with their hand extended over the Leap Motion Controller while wearing the Epson Moverio BT-100 Wearable Display attached with the PhidgetSpatial 3/3/3.

A within-participants design was used due to a low number of potential participants. This design also made it possible to compare users' performance when using different DPMs, and gather data on user preference based on their experience using all three DPMs. The order in which DPMs were tested varied between users in order to counterbalance any potential order effects.

Tasks & Procedures

The procedure of the user evaluation is as follows. First, the participant was given a brief introduction of the technology of the DPMs in the form a information sheet with concept diagrams of each DPM (see Figure 1's concept sketches). After providing information in a questionnaire about their individual use of computer desktop interfaces, demographic data, and hypothesis for each of our dependent variables, participants were put through the following DPM-specific procedure in a predefined order. Although only one DPM is referred, the procedure is executed once for each DPM – totalling 3 iterations.

At the start of this procedure, the participant was given a short explanation of the DPM and a maximum time of five minutes to warm-up. Each interface contained desktop objects; each object had a particular number and dominant colour. Once comfortable with the interface, the participant completed the following tasks while being prompted to narrate his/her thoughts aloud. Details of both tasks are described below. Participants completed the replication Task first so that they could to gain some sense of where the desktop objects would be placed for the retrieval Task.

1. *Replication Task*: The participant must replicate a 3D formation of desktop objects by dragging 15 icons to coloured sphere-areas that match the icon's main colour. Some icons were initially located within the incorrect coloured sphere, and they could be unintentionally dragged out of the sphere. An numerically-ordered list of the icons still needing to be moved was shown over the upper-left corner of the screen.
2. *Retrieval Task*: Now more aware of the location of certain desktop objects from completing the replication task, the participant was given the same interface with the icons grouped correctly where their coloured sphere was located, and asked to find and grab 15 desktop objects. Participants were anticipated to remember where the icons were placed by leveraging their spatial memory.

After the warm-up and each individual task, the participant was asked to give some general feedback and describe their experience in their own words. After the second task was completed, the participant then filled out a questionnaire to quantify the feedback, describe what they felt, and estimate their performance in all DPMs. This latter estimate was updated as they completed each DPM.

Once all DPMs were tested, the user was given a short questionnaire to re-evaluate their preference and experience with each DPM, and provide general suggestions or comments about their performance.

Measures

In this experiment, there are three dependent variables that were measured to determine the success of each projection model:

- *Spatial Memory Performance*: The total amount of time (in seconds) taken for a user to complete a single task. There is no time penalty for errors in either task, since users cannot complete the task without correcting the error.
- *Navigation Performance*: The physical path used to move to the destination, which is measured in terms of distance (in millimeters) the cursor was moved as shown in Equation 1.

$$\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \quad (1)$$

  This formula calculates the distance that the users' hands moved, Δx, Δy, and Δz are the hands' changes in the three dimensions. There is no penalty to distance for errors in either task, because users must correct the error in order to complete the task.
- *User Preference*: Subjective data on each participant's ease of use, difficulty at learning the each DPM, general preference, estimated effectiveness, and stress levels. These are measured using a questionnaire designed based on NASA-TLX [6].

Spatial memory and navigation performances from using each interface was compared in order to answer (1): if the iDPM will perform best (least time taken and shortest path taken). Comparing values of user preference will answer (2.): if the iDPM is most preferred by participants. In addition to these quantitative measures, we observed participants' comments, reasoning and common patterns among the gestures. This qualitative data is taken into consideration, along with user performance data, when analysing and discussing the results.

Data collection

The quantitative data was collected via computer-based automatic data recording and questionnaires. Data measuring user performance (Spatial Memory Performance, Navigation Performance) was collected automatically by the computer as users complete the tasks. Subjective data was collected using questionnaires. Within the questionnaire, users were asked to rate their different levels each aspect (e.g., ease of use) using a Likert scale (i.e., 1 = very low ease of use, 7 = very high ease of use).

Qualitative data was captured throughout the procedure. During tasks and in post-task feedback, narrations of the participant's reasoning and comments were recorded and transcribed for analysis. During the questionnaires, participants were prompted on why they selected different ratings. An interview with open-ended questions was conducted to collect some direct feedback on user preferences with the different DPMs in terms of how the DPM felt to learn and use, what was liked and disliked about it, and what suggestions could be made to improve it.

Pilot Study Results

After concluding this pilot trial, we re-assessed our preliminary study, and made some modifications to it: we removed the vDPM from quantitative comparison, and reduced the number of desktop objects in tasks. Our reasoning for changes is discussed below.

During the replication task for the vDPM, we found that it was almost impossible to use; the participant took more than 5 minutes just to grab the first desktop object. As mentioned in the previous section, the vDPM is similar to the pDPM, but with the addition of a third dimension. In order to grab a desktop object, the cursor (controlled by hand gesture) must overlap with the object in all three dimensions.Whenever a user tried to grab an object, the sensor (i.e., Leap Motion Controller) would interpret the grab gesture as a change in the user's hand position, causing the cursor to move away from the object, and making it very difficult for users to pick up desktop objects. Similar issues occurred when participants tried to release an object in a target area; the cursor would move out of the target area before releasing the object.

In the questionnaire conducted after using all three DPMs, the participant ranked the vDPM as the least preferred and yielding the worst overall experience. Therefore, we decided to have participants forgo the tasks with the vDPM, and instead just try the vDPM without measuring time and distance to provide only qualitative feedback.

Additionally, the number of icons in each task was reduced from 15 to 10. During the pilot study, the user spent more than 15 minutes completing the replication task in the iDPM, and needed rest breaks between tasks. This indicated that while using 15 icons, the study would take more time to complete than participants would be willing to spend on a voluntary basis. Ten icons were used because it was sufficiently many to still gather meaningful data on each DPM.

METHODOLOGY

In this experiment, we mainly evaluated the pDPM and iDPM. The goal of this study is to provide a more sound comparison of our two most promising DPMs. Because our pilot study determined the feasibility and effectiveness of our methodology, most of the procedure used in this experiment is nearly identical to the pilot. Only the following two modifications have been made to the experimental procedure:

1. The first change was to only qualitatively evaluate the vDPM. We let participants practice using the vDPM after completing the tasks with the other two DPMs. This change also affected our pre-defined order. Instead of having six different orders, we only had two, (i.e. pDPM → iDPM → vDPM, and iDPM → pDPM → vDPM). Time and distance for both tasks were only measured for the pDPM and iDPM..

2. The second change was to reduce the total number of desktop objects from 15 to 10 for both the replication and the retrieval task. This was done due to limitations with the amount of time available to conduct the study.

The experimental design and procedures were otherwise identical to that used in the pilot study as shown in Figure 3. New participants were used in this experiment, and they are described in detail in the next section.

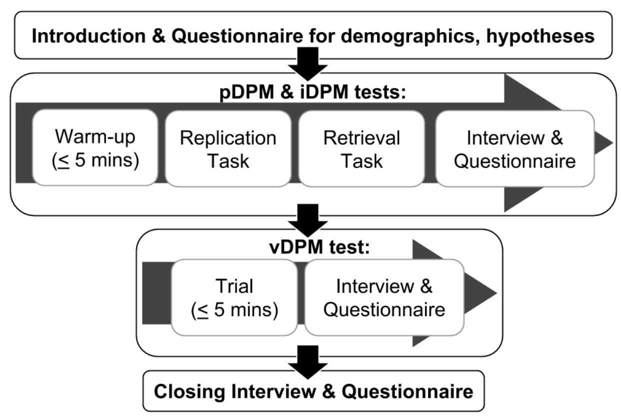


Figure 3. Experimental Design. In the experiment, participants were put through this procedure. To counterbalance order effects, planar (pDPM) and immersive (iDPM) tests were ordered differently between participants. The volumetric (vDPM) was conducted after these two, to gain only qualitative data.

Participants

In accordance to similar user evaluations [1, 8], this study used 6 new participants between the ages of 18 and 27. All participants were students at the University of Toronto – acquaintances of at least one of the researchers, recruited on a voluntary basis, and expected the study to involve a 3D interface and wearable display. All participants were required to have average control of their arms and hands, have either 20/20 vision or vision corrected to 20/20, and not be colorblind.

Participants, by filling in a questionnaire, reported that they used a desktop interface on their computers daily. Most (5 of 6) had never or very rarely used a gesture-based interface in the past year, while one had used one a few times a month. Half had used a drag-and-drop motion to move desktop objects sometimes, while the other half used it always. However, participants only used spatial memory to re-locate desktop icons usually, sometimes, and very rarely. The number of desktop icons varied greatly among participants: 1-9 (2 of 6), 10-24 (2 of 6), 25-49 (1 of 6), and 50-99 (1 of 6).

RESULTS

Based on our research questions, we found the following results relating to (1) the time and distance needed to complete the tasks and (2) user preference associated with each DPM.

Time & Distance

The average time and distance to complete each of the procedure tasks for the pDPM and iDPM interfaces are shown in Figure 4. Because the study used a within-participant design, and we wanted to inspect the performance measures of the pDPM and iDPM, we applied a Welch 1-sided paired t-test comparing each participant's performance in terms of time and distance for the separate replication and retrieval tasks. We confirmed using the Shapiro-Wilk test that these differences were found to be from a normal distribution. In all situations, the time and distance required when using the pDPM was significantly lower than when using the iDPM. The effect size of the DPM, according to Cohen's d, was large. The results of this study are presented in Table 1.

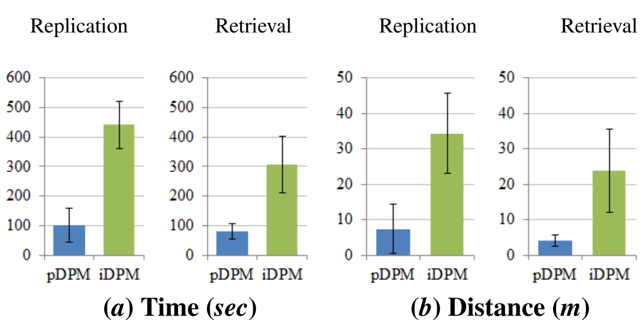


Figure 4. (*a*) Time and (*b*) Distance results of the pDPM and iDPM for each task (Replication or Retrieval). Bar heights indicate the average, while error bars indicate the standard deviation. In all cases, the pDPM's results are lower than the iDPM's results.

User Preference

The individual user preferences are shown in Figure 5. In the hypothesis (pre-tests) questionnaire, which asked users to rank their expected performance on each DPM based only on an introductory information sheet, most participants (5 of 6) expected that the iDPM would yield the best user experience. The pDPM was always perceived to yield the worst or medium experiences.

After participants had completed both tasks with all of the DPMs, most of them (4 of 6) rated the pDPM as being more preferred than the iDPM. By numerating this data (e.g, 3 = most preferred), and applying a Wilcoxon Signed-rank test,

Task	pDPM Mean	pDPM SD	iDPM Mean	iDPM SD	t(5)	d
Time (*sec*)						
1	100.77	56.69	440.37	79.44	10.19	4.16
2	80.33	24.48	306.40	95.07	5.78	2.36
Distance (*m*)						
1	7.40	6.85	34.32	11.37	8.79	3.59
2	4.05	1.67	23.85	11.73	4.22	1.72

Table 1. Paired t-test results comparing the planar (pDPM) and immersive (iDPM) DPMs using their required time (seconds) and distance (millimeters) to complete each tasks. Significance was found in all cases ($p < 0.05$). The effect size of the DPM, according to Cohen's *d* (*d*), was large.

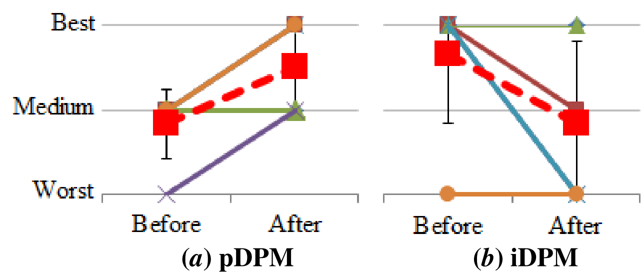


Figure 5. User Preference results for the pDPM and iDPM – as submitted by each participant before and after all tests in the procedure. The red dashed line represents the average rating (assuming Best = 3, Medium = 2, Worst = 1), and the error bars show the standard deviation.

we find no significance between the pDPM and the iDPM in both the hypothesis and conclusion ratings. The medians of each group and results of the study is show in Table 2.

To investigate further on how users ranked their preferences, we considered other ratings (e.g., mental demand, ease-of-use) and tried to find significant difference between the results for the pDPM and the iDPM.

DISCUSSION

The following sections aim to discuss the data gathered during the experiment, as well as the impact that different factors had on the data. We also discuss the issues with the vDPM that led to us not gathering quantitative data for it.

Time & Distance

In reviewing the data gathered on the time and distance required to complete each task, we consider various factors that could explain why the pDPM required less time and distance than the iDPM. This is of particular interest because our findings contradict H1; that spatial memory would allow users to

Ratings	pDPM Median	iDPM Median	Results W	Z	r
Pre-tests	2.0	3.0	18	1.67	0.48
Post-tests	2.5	2.0	6	1	0.29

Table 2. Wilcoxon Signed-rank test results comparing the planar (pDPM) and immersive (iDPM) DPMs using their given user preference (e.g., 3 = most preferred). In all cases, significance was not found ($p > 0.05$).

complete tasks more quickly using the iDPM. The data gathered also confirms H2; that the distance users need to move their hands is lower in the pDPM.

Planar DPM Factors
We believe that the familiarity that participants already have with a 2-dimensional desktop interface would have had a positive impact on their performance when using the pDPM. All participants in this study use a computer with a traditional desktop on a daily basis, and as a result, they were already familiar with the interface. The only learning that they were required to do was with the Leap Motion controller, which was a somewhat simple task.

Immersive DPM Factors
There were a number of factors that we believe had a detrimental effect on the performance of the participants while using the iDPM. Unlike other studies, such as Rädle et al. [14], which H1 is based upon, the iDPM has users moving over the surface of the sphere, rather than through the space inside it. Because the users work in a different space than they do in previous studies, we did expect some discrepancies between our results and the results in previous work.

The factor that we believe had the largest impact on user performance was the inability for a user to see their hands at all times when using this DPM. During the test, participants were required to keep their hand positioned over the Leap Motion controller, which was placed on a table in front of them. While using this DPM, participants would often be required to turn their heads and bodies, which often led to a situation where their hands were outside their field of view. When this occurred, the participants' hands would drift outside of the range of the Leap Motion controller, causing a navigation error which they then needed to correct. This issue artificially increased the time and distance needed to complete each task.

The next most significant factor is related to the first. Unlike the pDPM where users could see the entire desktop at all times, the iDPM limited the users to a 70° field of view, which often caused the participants to turn around to find specific icons or targets. Although this issue would not happen if the user was given more time to learn the position of each icon, it had a significant and negative impact during the experiment. This issue compounded with the first to cause a greater increase in time required for each task.

Another factor was the onset of dizziness in 2 of 6 of the test participants, one of whom regularly suffers from motion sickness. Unfortunately, there is insufficient data to suggest whether this was caused by the interface itself, or by the participants making several rapid head movements in succession. While there was some initial concern that the augmented reality glasses could cause nausea or motion sickness in some of the participants, none of the participants because ill or reported feelings of nausea.

Finally, we believe that the participants' expectations may have had an impact on their performance. Before the experiment, 4 of 6 participants expected the iDPM to require more time than the pDPM, and 2 of 6 expected it to take the same amount of time as the other DPMs. In terms of distance, all participants believed that the iDPM would require more physical distance during tasks. It's possible that the participants' expectations may have had a slight negative impact on their performance.

User Preference
In reviewing the user preference of each DPM, we take into account the levels of stress, self-aware performance, ease-of-learning, ease-of-use, mental demand and physical demand.

Ratings Prior to Experiments
Considering only the pre-tests preferences, most participants (5 of 6) complied with our hypothesis H3 that the iDPM would be the most preferred. Based on the concepts understood by the introductory information sheet, participants believed the iDPM would yield the best experience because it resembled popular science fiction, seemed "cool", and seemed to be the closest to reality. This last point coincides with the findings of Cockburn et al., which we based our hypothesis (H3) on, whose participants felt the 3D interface would be more "natural" [4].

Participants also generally rated the iDPM as having approximately the same ease-of-use as the pDPM. Hence, if the iDPM were to provide the same ease of use as the pDPM, we believe that it would be more interesting and subsequently more preferred. In contrast, the pDPM was usually rated as having a poorer experience because it did not seem as closely tied to reality as the iDPM (desktop objects were bounded to a 2D plane), or because it wasn't as unusual iDPM – suggesting some novelty component in initial preferences.

Ratings After Experiments
After the tests had been administered, most participants (4 of 6) preferred the pDPM over the iDPM. They cited their reasoning on implementation issues with the iDPM (e.g., the iDPM required users to maintain their hand position over the Leap Motion controller). The higher preference for the pDPM may have been because it proved better overall experience, ease-of-use, ease-of-learning, mental demand, physical demand, stress level, and self performance.

Considering the Likert-scale rankings provided for the overall experience (where 7 indicates the best experience, and 1, the worst), the pDPM and iDPM had medians of 6.0 and 3.0 respectively indicating that participants had a better experience in the pDPM. However, a Wilcoxon Signed-rank test found no significant effect of DPM on user experience ($W = 1$, $Z = -1.69$, $p > 0.05$, $r = 0.49$). Similar results were found in other measurements, as shown in Table 3. Hence, while the results suggest a more favourable experience with the pDPM, more trials may be needed to find significance.

There are two primary factors that may have influenced these measurements: influence from existing technology, and hardware limitations. First, users may have found the pDPM easier to learn and use due to its similarity to existing desktop interfaces. Second, there were some issues with the PhidgetSpatial that made it slightly more difficult to control the orientation of the camera when using the iDPM. This issue compounded with the limited range of the Leap Motion's

Type	pDPM Med.	iDPM Med.	W	Z	r
Ability to Learn	6.0	3.0	0	1.95	0.56
Ability to Use	6.0	3.5	0	2.11	0.86
Mental Demand	2.0	4.0	0	2.11	0.86
Physical Demand	2.0	4.0	0	1.95	0.79
Stress Level	3.0	6.0	17	1.37	0.39
Self-assessment of performance for each task:					
Replication	5.0	4.5	20	2.05	0.59
Retrieval	6.0	2.0	10	1.95	0.56

Table 3. Wilcoxon Signed-rank test results and medians (Med.) comparing the planar (pDPM) and immersive (iDPM) using their different properties of task load. In all cases, significance was not found ($p > 0.05$). These results that indicate that the pDPM was easier to learn and use, and had lower levels of mental demand, physical demand and stress. Participants also assessed themselves having done better for both tasks in the pDPM.

camera and made the iDPM much harder to interact with. These issues then increased the mental and physical demand required to use the iDPM and vDPM.

While the preference of the iDPM was generally worse than the pDPM, participants still seemed open the the its concept. Future work could overcome these implementation issues, and potentially produce more results that preferred the iDPM.

Volumetric DPM Feedback

In this section, we attempt to analyze some of the factors that caused users to be unable to use the vDPM. We will also analyze the participant feedback regarding the vDPM separately from the other feedback, since it was collected under different conditions; participants were not required to complete the tasks using the vDPM.

As mentioned as one of the issues of vDPM in the pilot study, one of the main factors was a limitation of the Leap Motion controller, which greatly increased the difficulty of interacting with the desktop icons. In order to “grab” an icon, users must close their first while the cursor is aligned with a desktop icon. Unfortunately, the algorithm used by the Leap Motion controller created an issue; the Leap Motion calculates the centre of a participant’s palm as the centroid of the hand. As a result, whenever users made a fist, the Leap Motion moved the palm position, causing the cursor to come out of alignment with the desktop icon.

The issue with the Leap Motion compounded with a second significant issue: the lack of depth perception. According to available documentation, it was impossible to use a stereoscopic display for this experiment. As a result, there was no depth perception in the vDPM. Additionally, although there were depth cues, users were not able to easily determine the relative positions of objects. The main consequence of this was that it became too difficult for users to move the cursor into alignment with desktop icons, making it much more difficult to grab them. This was compounded by problems found in Cockburn et al. [5], where it was shown that a 3D environment will create too much “wiggle room”, making it significantly harder to align the cursor with the icons.

As a result of these issues, participants could not be expected to complete either task; the participant in the pilot study took long enough to complete the first task that she asked to continue without completing it. Because of this, it was decided that asking participants to complete the tasks using the vDPM was unrealistic, and that it was more reasonable to simply ask the participants for qualitative data regarding this DPM.

Prior to testing, the participants’ predictions for the vDPM varied more than the other DPMs, with one-third of participants predicting that it would be the worst, average, or best DPM. After the experiment and understanding that researchers expected the vDPM to perform poorly, the participants’ opinions were similarly divided, with two participants thinking it was the best, one thinking it was average, and three thinking it was the worst.

As well, 2 of 6 users found the input mapping to be counterintuitive, and expected the mapping to be similar to a traditional desktop, where motion along the z-axis would move the cursor along the y axis. They commented that the input mapping felt unnatural and hard to use.

We believe that the divided opinions were a result of two conflicting factors. First, participants were not required to complete the tasks, which we believe lowered the mental stress of using it. As well, participants were able to use this DPM at their leisure, instead of trying to complete a task quickly, which would reduce the physical demand. These factors would have improved the participants’ experience. Opposing this factor is the increased difficulty of use that all participants experienced, as outlined above. Because of this conflicting factors, the participants’ opinions before and after the experiment were very similar.

Limitations

One of the issues that may have prevented users from being able to use the vDPM was that it was a 3D space being projected onto a 2D surface. As a result, users lacked the depth perception that they needed in order to properly use this DPM. This was due to technical limitations in the engine used to run the experiment, which did not support a stereoscopic display on this hardware. Another main limitations of this study is that the augmented reality glasses have a relatively small display (80” at 5m). As a result, it may not accurately reflect how well users could use each DPM when applied to a larger space, such as a 2’ display at a distance of 1m (a typical computer monitor).

CONCLUSION

In this paper, we presented a novel natural user interface to allow users to interact with desktop objects using augmented reality glasses and natural gestures, and attempted to improve performance and user experience through the use of spatial memory. In our study, we defined a Desktop Projection Model (DPM) as a way to represent a desktop in a 3D space, where we proposed three DPMs: planar (pDPM), immersive (iDPM), and volumetric (vDPM). To evaluate user performance, we collected and measured the spatial memory performance (time) and navigation performance (physical distance of gestures) based on two tasks designed to simulate typical desktop interactions.

From our pilot study, we found that vDPM was almost unusable, and hence, we only sought to obtain qualitative feedback for its design in our experiment. During our preliminary experiment, we found that the pDPM required significantly less time and less distance to complete tasks than the iDPM (answering research question (1) on user performance). Concerning user preference (2), we found that initially most people preferred the iDPM to the other DPMs; however, after use more people prefered pDPM. Finally, the vDPM was unusable due to hardware limitations.

The main limitations were from the augmented reality glasses and the Leap Motion controller. The augmented reality glasses used in our prototype had a relatively small perceived image size, i.e. 80" at 5 meters. The Leap Motion controller was the engineering prototype, which had a small effective range (approx. 1 meter), and a limited field of view (approx. 60 degrees). Thus, it was too sensitive for movement over the z-axis, and the hardware failed to detect hand position during grab gesture. We conclude that future work is necessary to determine the viability of this interface

Future Work

Due to issues with the vDPM, it would be beneficial to see how well this DPM can be used with a stereoscopic display. Our study looked at how well users can use each DPM when it's positioned relative to the user; in future, studies should also consider how users interact with each DPM when it has a static position (e.g., a fixed position in space), and users are able to move around it. For instance, the user could interact with the vDPM from the side, rather than from the front.

In this study, the sensor for tracking hand motion was positioned on a desk in front of the user. It may be possible to mount the sensor on the user's head, using either a harness of some sort or by attaching it to the augmented reality glasses. This may drastically increase usability, and should be investigated.

Finally, the effects of the iDPM on spatial memory must be more thoroughly and more explicitly evaluated (e.g., consider a greater time laspse between tasks). According to related research, the more body-centric interaction of the iDPM [14], and the additional difficulty of using it [3], would make it key in leveraging the spatial memory of a user – than the pDPM. Future work could analyse this potential effect, consider its implications, and design tasks more appropriate for the iDPM or the pDPM.

REFERENCES

1. Agarawala, A., and Balakrishnan, R. Keepin' it real: Pushing the desktop metaphor with physics, piles and the pen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, ACM (New York, NY, USA, 2006), 1283–1292.
2. Cockburn, A. Revisiting 2d vs 3d implications on spatial memory. In *Proceedings of the Fifth Conference on Australasian User Interface - Volume 28*, AUIC '04, Australian Computer Society, Inc. (Darlinghurst, Australia, Australia, 2004), 25–31.
3. Cockburn, A., Kristensson, P. O., Alexander, J., and Zhai, S. Hard lessons: Effort-inducing interfaces benefit spatial learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, ACM (New York, NY, USA, 2007), 1571–1580.
4. Cockburn, A., and McKenzie, B. 3d or not 3d?: Evaluating the effect of the third dimension in a document management system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, ACM (New York, NY, USA, 2001), 434–441.
5. Cockburn, A., and McKenzie, B. Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, ACM (New York, NY, USA, 2002), 203–210.
6. Hart, and Staveland. NASA Task Load Index. `http://www.test.org/doe/`, July 2006. Last Accessed on: March 17, 2014.
7. Jetter, H.-C., Leifert, S., Gerken, J., Schubert, S., and Reiterer, H. Does (multi-)touch aid users' spatial memory and navigation in 'panning' and in 'zooming & panning' uis? In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, ACM (New York, NY, USA, 2012), 83–90.
8. Lee, J., Olwal, A., Ishii, H., and Boulanger, C. Spacetop: Integrating 2d and spatial 3d interactions in a see-through desktop environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 189–192.
9. Malone, T. W. How do people organize their desks?: Implications for the design of office information systems. *ACM Trans. Inf. Syst. 1*, 1 (Jan. 1983), 99–112.
10. Mander, R., Salomon, G., and Wong, Y. Y. A "pile" metaphor for supporting casual organization of information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, ACM (New York, NY, USA, 1992), 627–634.
11. Mann, S. *Intelligent Image Processing*. John Wiley & Sons, Inc., New York, NY, USA, 2001.
12. Mistry, P., and Maes, P. Sixthsense: A wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches*, SIGGRAPH ASIA '09, ACM (New York, NY, USA, 2009), 11:1–11:1.
13. Mistry, P., Maes, P., and Chang, L. Wuw - wear ur world: A wearable gestural interface. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, ACM (New York, NY, USA, 2009), 4111–4116.

14. Rädle, R., Jetter, H.-C., Butscher, S., and Reiterer, H. The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, ACM (New York, NY, USA, 2013), 23–32.

15. Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., and van Dantzich, M. Data mountain: Using spatial memory for document management. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology*, UIST '98, ACM (New York, NY, USA, 1998), 153–162.

16. Robertson, G., van Dantzich, M., Robbins, D., Czerwinski, M., Hinckley, K., Risden, K., Thiel, D., and Gorokhovsky, V. The task gallery: A 3d window manager. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, ACM (New York, NY, USA, 2000), 494–501.

17. Scarr, J., Cockburn, A., Gutwin, C., and Malacria, S. Testing the robustness and performance of spatially consistent interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 3139–3148.

18. Whittaker, S., and Hirschberg, J. The character, value, and management of personal paper archives. *ACM Trans. Comput.-Hum. Interact. 8*, 2 (June 2001), 150–170.

19. Wigdor, D., and Wixon, D. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*, 1st ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011.