

Exploring 3D Interaction in Alternate Control-Display Space Mappings

Jeroen Keijser* Sheelagh Carpendale* Mark Hancock* Tobias Isenberg*
Department of Computer Science
University of Calgary, Canada

ABSTRACT

The desire to have intuitive, seamless 3D interaction fuels research exploration into new approaches to 3D interaction. However, within these explorations we continue to rely on Brunelleschi’s perspective for display and map the interactive control space directly into it without much thought on the effect that this default mapping has. In contrast, there are many possibilities for creating 3D interaction spaces, thus making it important to run user studies to examine these possibilities. Options in mapping the control space to the display space for 3D interaction have previously focused on the manipulation of control-display ratio or gain. In this paper, we present a conceptual framework that provides a more general control-display description that includes mappings for flip, rotation, skew, as well as scale(gain). We conduct a user study to explore 3D selection and manipulation tasks in three of these different mappings in comparison to the commonly used mapping (perspective mapping of control space to a perspective display). Our results show interesting differences between interactions and user preferences in these mappings and indicate that all may be considered viable alternatives. Together this framework and study open the door to further exploration of 3D interaction variations.

Keywords: 3D interaction, control-display space mappings, evaluation.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/Methodology; I.3.4 [Computer Graphics]: Graphics Utilities—Virtual Device Interfaces; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques

1 INTRODUCTION

As a culture, we have had fantastical ideas about exploring alternate 3D spaces for at least a century – take for example “Through the Looking Glass and What Alice Found There” by Lewis Carroll [7]. Currently, a great variety of possible 3D interfaces — Caves™, 3D visualizations, 3D games — continue to proliferate. These interfaces differ in many aspects, including purpose (work, play, or entertainment), display capabilities and configurations, and input devices. Most 3D interfaces make use of the familiar 3D mapping derived from Brunelleschi’s perspective that has been used in computer graphics for the last forty years or more. There are some exceptions, such as various layered 2D interfaces to provide fast 3D effects in games, and recently there has been some exploration into the creation of 3D scenes in which multiple perspective variations are integrated [1, 8]. Embellishments are being explored, such as passive or active stereo and 3D sound effects, which can add to the power of the 3D experience, however, the basic 3D perspective commonly stays the same.

There is significant amount of research into different aspects of 3D interfaces, such as 3D displays (e.g., [11, 21]), 3D input devices (e.g., [25]), and 3D interaction metaphors (e.g., [16]). In

parallel there is also a continuing discussion about difficulties with 3D interfaces including problems with developing effective 3D perceptual support (e.g., [4, 22]), problems with developing intuitive and understandable 3D visualizations [18], and more generally discussions about lack of overall adoption of 3D interfaces [2]. One aspect that has received less research attention is the mapping of the interaction control space to the display space. To explore alternate experiences in 3D interaction, we consider the 3D to 3D mapping from a person’s interaction space in the real 3D world (*control space*) onto the virtual 3D *display space*. While control to display space mappings have been explored in terms of relative scale (gain) [12, 20, 22], there has been little exploration of whether there are other viable alternatives. Despite other differences amongst 3D interfaces, most explorations into 3D interaction rely primarily on perspective for the creation of the 3D display space and commonly map the interactive control space directly into this perspective display space. While this mapping from the interactive control space to the display space appears to be an obvious choice, it does cause some perceptual discontinuities. For example, if one moves the input in control space to the lower left hand corner just in front of the screen, one’s cursor in virtual 3D display space is actually considerably to the right and higher up the screen (see Figure 1).

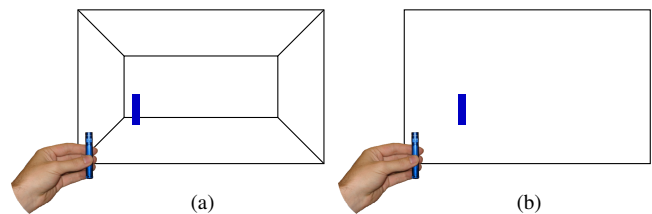


Figure 1: Mapping control space to display space: (a) shows a perspective display space with a perspective control/display mapping for the interaction. Note how the corresponding cursor object in display space makes sense within the box drawn in perspective but when the box cues are removed (b) this mapping is not so obvious.

These perceptual discontinuities raise questions about whether this style of 3D interaction is understandable due to our ability to comprehend perspective depth cues, or whether some alternate control to display space mapping would actually be easier to work with. We explored many alternative mappings and developed a conceptual framework to describe them. When considering which mappings to explore in depth, we looked for metaphors within our everyday world that might lead to alternate but comprehensible control to display space mappings. One choice is to address the issue drawn in Figure 1 by straightening out the distortion created by the common perspective mapping – an orthographic interaction in which the display is still in normal perspective but the control to display mapping is one-on-one for *x*- and *y*-movements. Our second choice is to explore a ‘through the looking glass’ metaphor (Figure 2), which would draw on culture history from “Through the Looking Glass and What Alice Found There” [7] and with our experiences with mirrors in our everyday lives. We conducted a study that explores perspective, orthographic and mirrored control to display space mappings for 3D interaction.

*e-mail: {keijser|sheelagh|msh|isenberg}@cpsc.ucalgary.ca

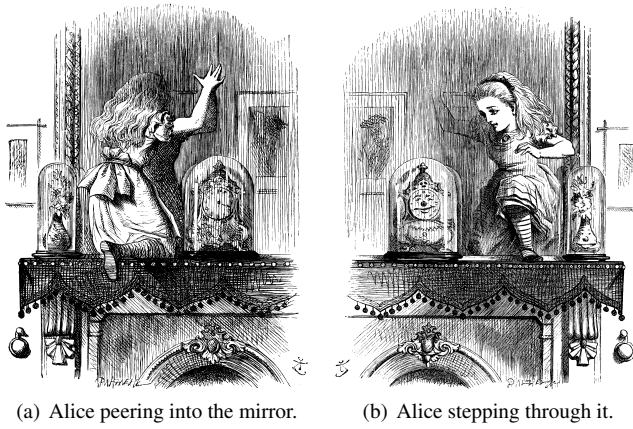


Figure 2: ‘Through the looking glass’ metaphor, from [7].

In this paper, we present our conceptual framework that provides a more general control-display description that includes mappings for flip, rotation, skew, as well as scale (gain). We conducted a user study to explore 3D selection and manipulation tasks in three new mappings in comparison to the commonly used mapping (perspective mapping of control space to a perspective display). Our results show interesting differences between interactions and user preferences in these mappings and indicate that all may be considered viable alternatives. Together our conceptual framework for control to display space mapping and our study of these interaction spaces open the door to further exploration of 3D interaction variations.

In the following section we discuss the related research and then in Section 3 we describe our conceptual framework. Section 4 presents our study, followed by results in Section 5 and discussion in Section 6. Section 7 draws our conclusions.

2 RELATED WORK

In the literature there is a wealth of knowledge and research into the effects of input choice and design (e.g., [25]), the effects of display choice (e.g., [11]), and the perceptual cues displays support for 3D interaction (e.g., [4]). However, the effect of the relationship between input or control space and output or display space has not been investigated in a comprehensive way.

There are many components to this relationship and we discuss them in Section 3, but one component of the mapping relationship that is understood, studied, and manipulated extensively is *scale*. Referred to as control/display ratio [3] or control gain [12], scale can be dynamically changed to extend direct manipulation (go-go gadget technique) or to dynamically change navigation control such as for a Wand in a CAVE™ or for more precise manipulation as in the PRISM technique [9]. Scale has also been used to create multiple views at different scales such as in World In Miniature (WIM) [20]. WIM is a technique to allow a user to interact at two different scales, one locally and the other in the global overview as represented by the WIM. Semantic Pointing [3] suggests that the control/display ratio should be adjusted to allow more control around semantically important components in an interface. Although less than positional scale, rotational scale has also been looked at. Poupyrev et al. [17] studied rotational control gain as part of their study into non-isomorphic rotational mappings. In addition to this selection of techniques there exists a multitude of other interaction techniques and studies that use or address scale.

Other studies that focused on how well humans can use current direct manipulation systems, but did not look directly at the control to display mapping, noticed some discrepancies between movement in the three main axes. Zhai et al. [26], for example, noticed performance differences. As expected, performance in the depth di-

mension (z) was the worst, distances in the vertical dimension (y) were overestimated, and the horizontal dimension (x) was the most accurate. Grossman and Balakrishnan’s 3D Fitts’ Law study [10] showed that the a target size should be larger in depth dimension (z) and that drastic size differences between dimensions (more than a factor of 4) are not helpful in increasing ease of target selection.

In order to successfully support the mapping from control to display space it is also necessary to understand the cognitive basis of the human comprehension of three-dimensionality. This is an ongoing area of research in psychology with much still unknown, however, one hypothesis indicates that humans have two models of the world around us: a cognitive model and sensory-motor model [13]. The cognitive model is world-centric and based mostly on our visual system which uses a wide variety of depth cues, some of which are selectively based on the task we are doing (needle vs. driving) [23]. The sensory-motor model or kinesthetic understanding of the world, in contrast, is egocentric and is based on a sensory-motor implicit understanding of the location of our body in relation to the world. It is mainly influenced through touch or the haptic channel. Burns et al. [6] note that people are more aware of the cognitive model and are generally adaptable to conflicts between the cognitive and sensory-motor models.

In 3DUIs, due to the inherent disconnect between control space and virtual space, and due to the state of current hardware, some sensory cues are comprised. Virtual objects cannot simply be picked up and manipulated using our haptic channel. With the exception of volumetric displays, the display of 3D is done on a 2D surface. Therefore, the depth dimension is artificially constructed rather than inherently supported. Perspective projection is the most common depth cue used to support this, and underlies most 3DUIs. The relative benefits of this and other depth cues have received considerable research attention. For instance, stereoscopic vision improves understanding of 3D scenes, but due to the difference in accommodation and convergence it causes eye strain [11]. See Boritz [4] for more on the differences between stereopsis and motion parallax as well as Wanger et al. [23] for other depth cues.

3 A CONCEPTUAL FRAMEWORK FOR CONTROL TO DISPLAY SPACE MAPPING

In order to be able to discuss the specific mappings we examined in a consistent way we first introduce a general framework that can be used to model and understand such mappings between control and display space. We start by discussing and analyzing scale, order, lag, orientation, flip, and skew as alternate mappings, then focus on the flip and skew mappings used in our study.

3.1 Control-Display Space Mappings

A 3D control space can be mapped to a 3D display space in many different ways. In general, this control-display space mapping can be described as a function, f , that maps the change in input degrees of freedom (DOF) in control space to a change in position and/or orientation in the display space:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (1)$$

where n is the DOF of the control space and m is the DOF of the display space. We specifically describe the use of a 3DOF input device for controlling the 3D position of an object, and so we specifically consider mappings of the form:

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad (2)$$

As detailed below, many different functions could prove useful for the control of an object’s position in 3D.

The control/display ratio [3] or control gain [12] defines the *scale* or size relationship between control and display space. Scale functions are of the form:

$$f(\Delta x, \Delta y, \Delta z) = (s_x \Delta x, s_y \Delta y, s_z \Delta z), \quad s_i \in \mathbb{R} \quad (3)$$

Although it is possible to have different scales for each axis and angle of rotation, typically the same scale is used for all axes ($s_x = s_y = s_z$). The effect of scale as well as its manipulation and has been used previously in a number of systems. As discussed in Section 2, examples include techniques such as World in Miniature (WIM) [20] which allows a user to interact at two different scales, the Go-Go technique, and other scaled metaphors which adjust scale to increase range of direct manipulation. Wang and MacKenzie [22] studied the relative scale of input device, cursor, and target size and noted that a direct one-to-one scale was easiest to use ($s_i = 1$).

It is also possible to differentiate between position-, rate-, and acceleration-based control of the display space. Poulton [15] describes this difference as a change in *order*. For example, a rate-based function could be the following:

$$f(\Delta x, \Delta y, \Delta z) = \left(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t} \right) \quad (4)$$

Zero-order or position control maps the control directly to position in virtual space, first-order or rate control maps the control to the velocity of movement in display space, and second-order or acceleration control maps the control to the acceleration of movement in display space. Position control is the most common. Velocity control is usually used with an isometric control such as a joystick. Acceleration control is normally not done in practice, but it may have special applications worth considering.

Lag can occur either because of deficiencies in the input device or can be introduced into the control-display space mapping. Such a function would require storage of past input data obtained. One possible lag function could be as follows:

$$\begin{aligned} g_t(\Delta x, \Delta y, \Delta z) &= (\Delta x, \Delta y, \Delta z) \\ f_t(\Delta x, \Delta y, \Delta z) &= g_{t-L}(\Delta x, \Delta y, \Delta z), \exists L \in \mathbb{R} \end{aligned} \quad (5)$$

Lag is normally thought of as detrimental to interactivity, however in some cases deliberate delay may improve performance [14].

Orientation refers to the orientation of the control space with respect to the display space. The entire space can be rotated by an angle around one of the coordinate axes or an arbitrary axis. For example, a rotation about the z -axis by an angle of θ would have the following control-display function:

$$f(\Delta x, \Delta y, \Delta z) = (\Delta x \cdot \cos\theta + \Delta y \cdot \sin\theta, \Delta y \cdot \cos\theta - \Delta x \cdot \sin\theta, \Delta z) \quad (6)$$

The effect of orientation has been studied in two dimensions by Wigdor et al. [24], but the 3D case has not been addressed to date.

The mapping can also *flip* one of the coordinate axes. Technically, this can be considered as a subset of scaling, however, flipping an axis may affect the user's understanding in a profoundly different way than would scaling. For example, a flipped depth axis ($s_z = -1$) would cause the interaction to be like that in a mirror.

Finally, a *skew* mapping is rarely considered. For example, analogous to perspective projection of world coordinates to display coordinates, the mapping from control to display space may also be subjected to a perspective distortion, thus spatially skewing the mapping. The control-display mapping function would be:

$$f(\Delta x, \Delta y, \Delta z) = M \cdot \Delta \vec{x} \quad (7)$$

where M is the projection matrix and $\Delta \vec{x} = [\Delta x \ \Delta y \ \Delta z \ 1]^T$. In typical 3D applications, the control space is mapped using the same projection matrix as is used for mapping the world to the display. However, by applying an inverse distortion to the mapping we can control the degree of the effective distortion. To our knowledge, these last two mapping types, flip and skew, have not been discussed in the literature to date.

3.2 Our Study's Four Selected Mappings

In this section we discuss the effects of flip and skew in detail and derive the four mappings we examined in our study as special cases and combinations of these types.

The perspective projection often used in rendering by default as a visual effect also results in a distortion of the control to display space mapping. This distortion causes forward and back motions in physical space to be mapped to perspective lines in virtual space. This agrees with the visual representation of the depth parameter. However, perspective lines move towards the center of the screen, and as such the cursor has a sideways motion while in control space the motion is perpendicular to the display surface. This disagrees with the kinesthetic representation which requires that forward motion in real space is mapped to a motion directly or orthographically forward in virtual space. We call the former mapping without correction of the perspective projection distortion *visual concordance*; and the latter mapping that uses an additional inverse perspective distortion to eliminate the effects of perspective projection for the interaction mapping *kinesthetic concordance*.

In addition to skew mappings, flip mappings may be relevant to 3D interaction. In particular the flipping or mirroring of the z -axis seems to be promising since we are familiar with mirrors in the real world. Thus, it seems natural to examine control to display space mappings that mirror depth such that a motion toward the screen in control space is mapped to a mirrored cursor motion toward the front of the scene in display space. This mapping can then be compared to the more common mapping that translates a motion toward the screen in control space to a cursor motion away from the display surface. This difference may be relevant for interaction with large displays where users are able to get close to the screen and could take advantage of interacting with objects directly at the display surface that the mirrored mapping facilitates.

The four mappings below are derived from varying the two conditions of each of these two parameters. Perspective mapping is the de facto standard mapping. Mirrored Perspective mapping is the mirrored version of the standard mapping where the depth axis is flipped. Both of these agree with visual concordance. Orthographic mapping and Mirrored Orthographic mapping are the result of eliminating the mapping distortion caused by perspective projection and its mirrored version, both represent kinesthetic concordance.

3.2.1 Perspective and Mirrored Perspective Mapping

Figure 3(a) shows a schematic view of the standard mapping from interactive control space to virtual display space that results from perspective projection. The interactive control space is shown below the thick horizontal line representing the screen; the virtual 3D space is above this line. The top quadratic box corresponds to the clipped viewing volume, the angled dark blue dashed shape represents a perspectively distorted box that is used as a reference point in the virtual space. The dashed lines in control space refer to the reverse (mental) mapping of this box back into interactive space. The letters are used to indicate how specific points are mapped from control to display space due to skew and flip differences between the mappings. Notice that points in interactive space shift position and are closer together at large depth inside the perspective box.

Figure 3(b) shows the standard mapping mirrored in z -direction. Notice that the points in interactive space are still all inside the perspective box as above, but that symbols and positions are now mirrored. A point near the screen in real space is now close to the screen in virtual space.

3.2.2 Orthographic and Mirrored Orthographic Mapping

Figure 4(a) demonstrates that through an additional inverse perspective distortion the effect of the perspective projection is removed, resulting in an undistorted mapping from control to display space.

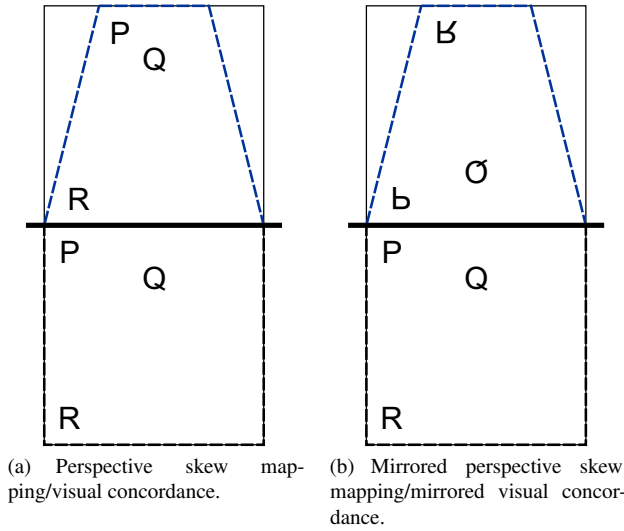


Figure 3: Schematic illustration of the perspective and mirrored perspective mappings. Below the thick horizontal lines representing the display surface are the interactive control spaces, above the line are the virtual display spaces. The thick dashed lines represent the space users are moving in, both for control space (black) and display space (dark blue). The dashed dark blue lines in display space also represent the perspective projection of a box. The letters indicate positions of points in control space and the mapped positions of these points in display space.

The points in interactive space are now mapped directly or orthographically into virtual space. Notice that motions perpendicular in control space (from R to P) now also map to motions perpendicular to the display surface in display space, matching the kinesthetic model. This has the effect that the virtual cursor is always directly in front of the physical cursor. Finally notice that the perspective reference box in virtual space now also maps back to a perspective box in interactive space.

In Figure 4(b) mirroring is applied to the z-axis, creating a mirrored orthographic mapping. Notice that motion paths perpendicular to the display surface in control space still map to motion paths perpendicular to the display surface in display space. However, since they are mirrored at the screen surface, points close to the screen in interactive space are close to the screen in display space.

4 STUDY DESIGN AND METHOD

Although the framework applies to 3D interaction in general, for the purposes of the study we limited ourselves to looking at selection within direct manipulation. Our tasks (see Section 4.3) reflect this and are based upon common study selection tasks, namely a point or volume location target [4, 5, 10]. Fitt's Law is a well accepted model of point or area selection, has been studied extensively in 2D, and has been extended to 3D [10]. In the tradition of Fitt's Law studies we based two of our tasks on the 2D multi-directional tapping task described in ISO Standard 9241 Part 9 [19].

As we wish to investigate people's performance as a result of differences in skew and flip, all other parameters remain constant. Since mirrored (flipped) mapping is of interest, in particular, when users interact close to the surface of the screen, the distance is kept to zero. The influence of the time parameters is kept as minimal as possible and results from the lag of the tracker. We use a scale factor of 1 for all axes and, as mentioned before, use a position-based control.

We used a 2 (*flip*: mirrored, non-mirrored) \times 2 (*skew*: perspective, orthographic) within-subjects design resulting in a total of four

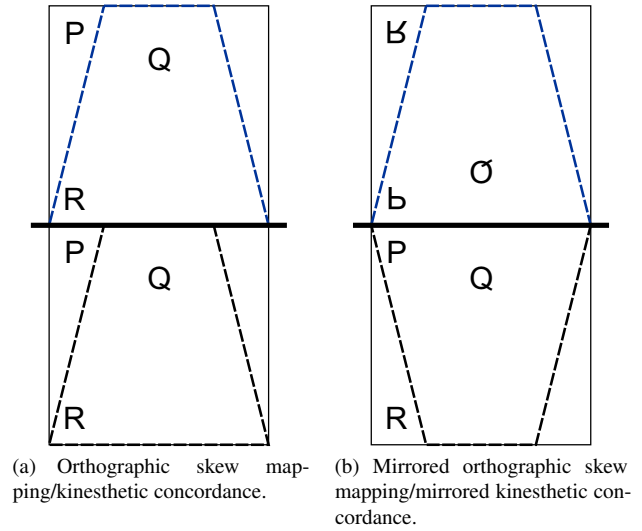


Figure 4: Schematic illustration of the orthographic and mirrored orthographic mapping. Notation as in Figure 3.

mappings.

4.1 Participants

Twenty-four people (13 male, 11 female) participated in our study, recruited through e-mail, notice of the day, and word of mouth. Four of the participants were left-handed and the remaining 20 were right-handed. 15 participants reported having little to no previous 3D gaming/graphics experience and nine reported having moderate to a lot of previous 3D gaming/graphics experience.

4.2 Apparatus

The display used for all four tasks was a 1024 \times 768 pixels, 73.3 cm \times 55.0 cm wall display. In all cases, the display space was a perspective grid representing five walls of a virtual room with lighting as an additional depth cue (see Figure 5). Participants were able to control a 3D cursor using a tracked light pen in the 73.3 cm \times 55.0 cm \times 55.0 cm volume directly in front of the display. Participants were given the option to sit, but all chose to stand throughout all trials. The pen was tracked using vision algorithms and input from two cameras mounted directly above the control space. Infrared filters were used to detect the near infrared light from the pen light (non-LED). Two computers were used in the experiment: one to capture movement via the two cameras that was then sent to the second (main) computer which controlled the display environment. Movement in the control space was mapped (using the mappings discussed above) to the movement of a cursor in the display space. The cursor was represented at a one-to-one scale (as suggested by Wang et al. [22]) in the display space as a cylinder texture-mapped with an image of the light pen. Crosshairs have been used previously in 3D selection studies [10] and we use crosshairs extending to each wall of the display space as additional feedback for the position of the cursor. Thus, the input used was an absolute indirect untethered 3DOF tracking device.

4.3 Procedure

For each mapping, participants performed four tasks in the same order, followed by a break task (see Table 1). The order of mappings was different for every participant, so that all 24 possible orderings were each used exactly once. Before each mapping was used, the experimenter demonstrated the mapping to the participant.

First Task—Grid Task: For the Grid Task each participant was asked to tap a sequence of 13 orange squares on the checkerboard

Step	Mapping	Task
1	1	Grid
2	1	Static ISO
3	1	Moving ISO
4	1	Photo Box
5		Break Task
6	2	Grid
7	2	Static ISO
8	2	Moving ISO
9	2	Photo Box
10		Break Task
11	3	Grid
12	3	Static ISO
13	3	Moving ISO
14	3	Photo Box
15		Break Task
16	4	Grid
17	4	Static ISO
18	4	Moving ISO
19	4	Photo Box
20		Break Task

Table 1: The order in which tasks were done by participants.

walls with the bottom of the light cursor. A single orange square would appear one at a time in a grid section of the alternating gray grid on the walls of the room (see Figure 5).

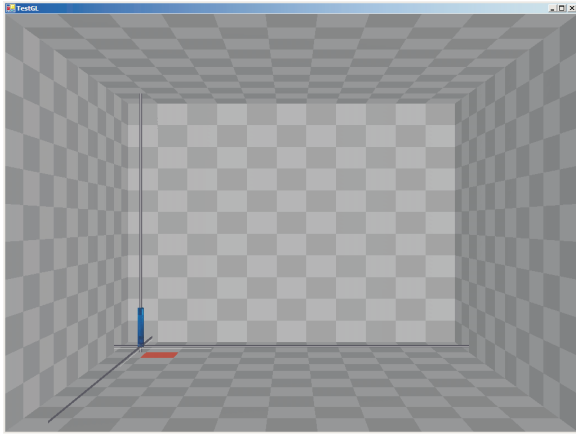


Figure 5: Grid Task. Participants were asked to tap orange squares that appeared on the checkerboard walls with the bottom of the blue light cursor.

Second Task—Static ISO Standard Task: Inspired by the multi-directional tapping task described in the ISO Standard 9241 Part 9 [19], participants were asked to tap the top center of cylinders placed on a circle. To indicate which cylinder was next, the cylinder would pop up to a random height with a crosshair on the top (see Figure 6).

Third Task—Moving ISO Standard Task: This task is similar to the previous task, except, instead of a pillar placed at a random height, it rises slowly from the floor, creating a moving target.

Fourth Task—Photo Box Task: Participants were asked to sort a set of cubes, randomly placed on the bottom of the virtual room, with images on the sides into what they thought was either an inorganic or an organic image (see Figure 7). They would do this by picking up the cube and placing it either on a blue mat or green mat at the front of the virtual room.

Break Task—Sketch Task: This task was intended as a break between sets of tasks, in order to rest the arm of the participant,

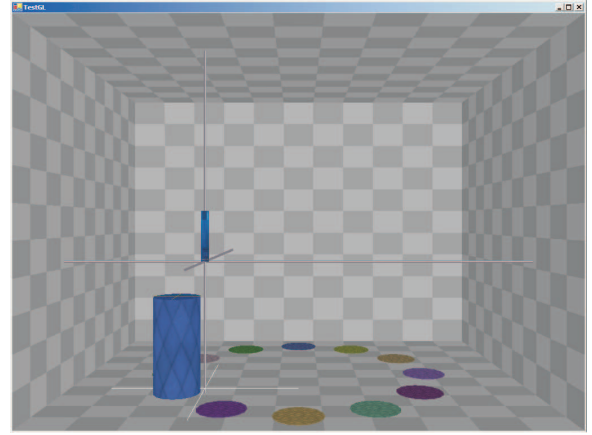


Figure 6: Static or Moving ISO Task. Participants were asked to tap the top of the raised cylinder with the bottom of the light cursor. The cylinders that were to be tapped were either not moving (Static ISO Task) or slowly raising from the ground (Moving ISO Task).

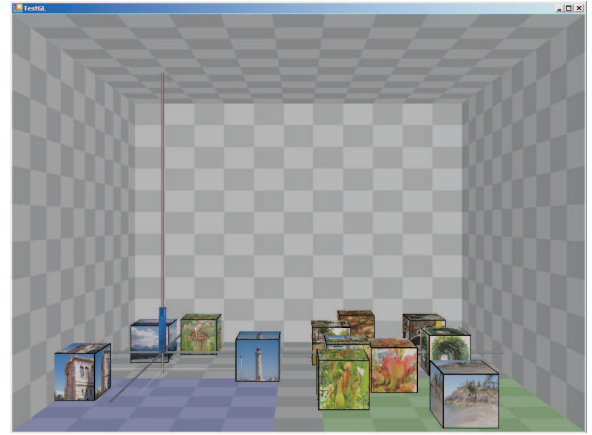


Figure 7: Photo box task. Participants were asked to sort the textured cubes into those showing inorganic or organic images and according to this decision place them on either of the two colored mats at the front of the virtual room.

as well as give feedback on the top down diagrams of mapping spaces. Participants were asked to draw the location of two points in interaction space in the virtual space, after the experimenter drew out and explained what we understand as the interaction space, the screen, the virtual room, and the space behind the walls.

4.4 Data Collection

Each session was video recorded to capture the motion of the user and audio comments from the participants as they were using the system. The system logged the virtual tracker motion, as well as any events that occurred (e. g., picking up a box or tapping a cylinder).

5 RESULTS

We performed a $2(\text{skew}) \times 2(\text{flip})$ within-subjects factorial ANOVA on the task completion times for the grid task and both ISO tasks.

5.1 Grid Task

There was a significant main effect of flip in the grid task ($F_{1,23} = 13.1, p = .001$) and a marginally significant effect of skew ($F_{1,23} = 3.3, p = .08$). The interaction between skew and flip was not significant ($F_{1,23} = 1.1, p = .30$). Participants performed the task

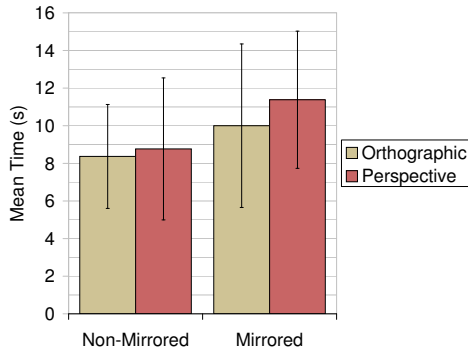


Figure 8: Grid Task: mean task completion times for the grid task. The mirrored mappings were slower than the non-mirrored and the orthographic skew mappings were faster than the perspective skew mappings. There was no interaction between flip and skew.

more quickly when the mapping was not mirrored ($M = 8.57$, $SD = 3.28$) than when it was ($M = 10.69$, $SD = 4.03$). Participants performed the grid task more quickly in the orthographic mappings ($M = 9.18$, $SD = 3.70$) than in the perspective mappings ($M = 10.07$, $SD = 3.90$). Figure 8 shows the mean task completion times for each mapping.

5.2 Static ISO Task

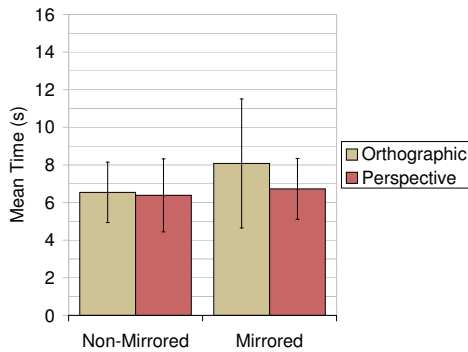


Figure 9: Static ISO Task: mean task completion times for the static ISO task. The mirrored mappings were slower than the non-mirrored mappings, and the perspective skew was faster than the orthographic skew in the mirrored case.

For the static ISO task, there were significant main effects of flip ($F_{1,23} = 5.5$, $p = .03$) and skew ($F_{1,23} = 6.0$, $p = .02$) and a marginally significant interaction between flip and skew ($F_{1,23} = 3.4$, $p = .08$). Post-hoc analysis revealed that, in the non-mirrored case, the difference between orthographic and perspective skews was not significant ($p = .68$), but for the mirrored case, participants performed faster with the perspective skew than with the orthographic skew ($p = .01$). Figure 9 shows the mean task completion times for each mapping.

5.3 Moving ISO Task

There was a significant main effect of flip in the moving ISO task ($F_{1,23} = 15.5$, $p = 0.001$). Participants completed the task more quickly when using a non-mirrored mapping ($M = 6.28$, $SD = 1.58$) than when using a mirrored mapping ($M = 7.96$, $SD = 2.72$). There was no main effect of skew ($F_{1,23} = 0.1$, $p = .76$) and no interaction between skew and flip ($F_{1,23} = 0.4$, $p = .54$).

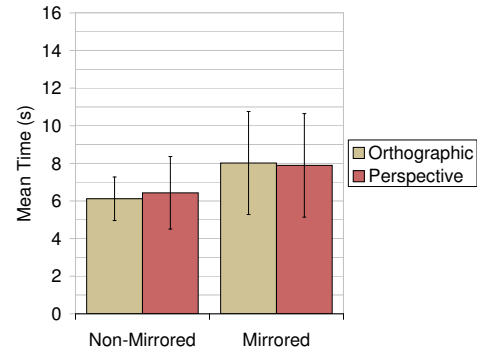


Figure 10: Moving ISO Task: mean task completion times for the moving ISO task. The mirrored flip mappings were slower than the non-mirrored mappings, and there was no main effect of skew.

5.4 Subjective Preference

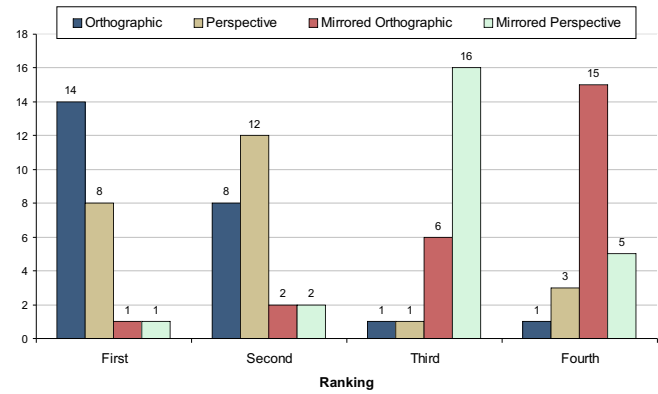


Figure 11: Overall mapping preference ranking: participants were asked to rank mappings from most preferred (first) to least preferred (last).

After completing all tasks in all four mappings we asked participants to rank the four mappings in order of preference (see Figure 11). Participants showed a clear preference for the non-mirrored mappings (92% first, 83% second, 8% third, 17% fourth). Most participants chose the orthographic non-mirrored mapping as their first choice (58%) followed by the perspective non-mirrored (33%) and the two mirrored mappings were each chosen as first by only one participant.

5.5 Trajectory Traces

From our log data we reconstructed and traced the path that participants took between target points. Typically, these trajectories show that participants generally had little difficulty going from one target to the next and took an almost direct path to get there. However, in every mapping almost all participants would at some point have difficulty tapping a target. Some participants were consistently more direct in both non-mirrored mappings, others in both orthographic mappings, and still others in both perspective mappings. However, even in the mappings a particular participant was most direct in, he/she would have difficulties at random times and for all tasks.

In each of the trajectory trace Figures 12 to 15 (based on example data of difficulty in each mapping) two orthogonally rendered views are shown, one from the front and one from the side of the trajectory a participant took between two targets. The orthogonal

rendering allows us to look at x - y -projection in the front view and y - z -projection in the side view. The target is circled in red.

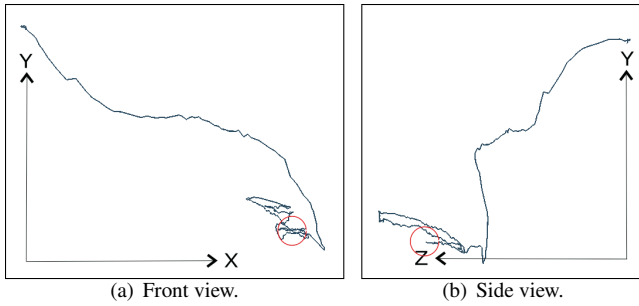


Figure 12: Trace between two targets in orthographic mapping. Notice the amount of motion around the target particularly in z -direction.

In Figure 12 an example of difficulty in the non-mirrored orthographic mapping is shown. The participant quickly and directly goes to approximately the right location and then has difficulty in the depth dimension z and some difficulty in the horizontal dimension x .

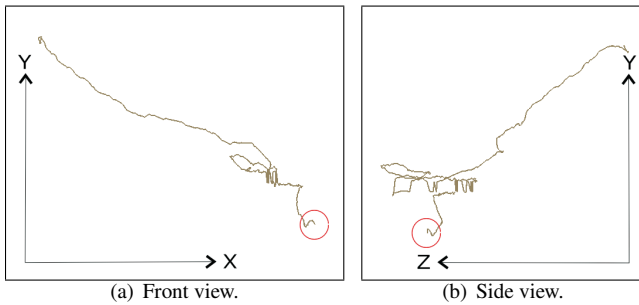


Figure 13: Trace between two targets in perspective mapping. Notice that the user becomes lost in depth slightly above the target.

In the example trajectory in Figure 13 of the non-mirrored perspective mapping, a participant gets close to the target and then has a lot of movement in z and y just above the target before they can orientate themselves again and move directly down to the target.

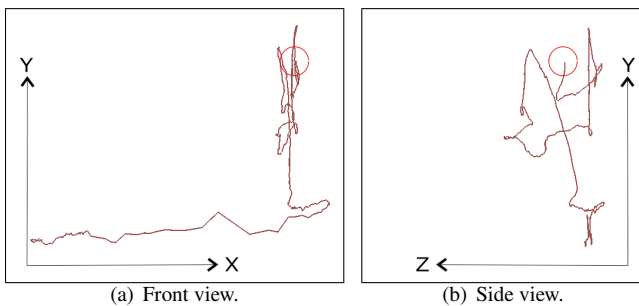


Figure 14: Trace between two targets in mirrored orthographic mapping. Notice the amount of motion in y - and z -direction.

In the mirrored cases Figures 14 and 15, from the front view it appears the participant goes almost directly there with some adjustment in y , but in the side views one can notice that the participant is in fact having much difficulty with the depth dimension z .

5.6 Quotes

Participants had a wide variety of responses to the differences in mappings. Generally as reflected in the ratings above (see Sec-

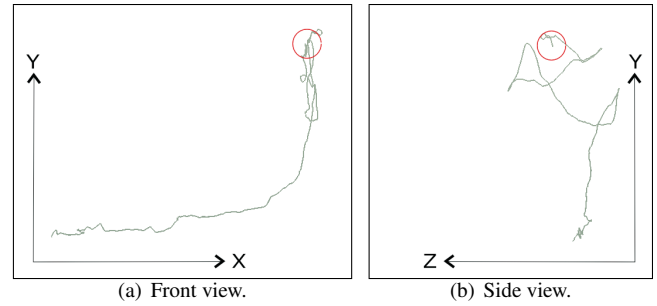


Figure 15: Trace between two targets in mirrored perspective mapping. Notice the depth confusion.

tion 5.4), mirrored mappings gave participants the most difficulty: “The A and B [non-mirrored] seemed more user-friendly, and the C and D [mirrored] seemed counteractive to my visual way of spatial sorting.” Some participants therefore strongly disliked it, “Damned mirrored!,” others enjoyed this challenge, “kinda fun because challenging” (mirrored perspective), “I liked the other [non-mirrored] method better but this one [mirrored] I got more of a sense of accomplishment once I grasped it!” However, a few users in fact did better in and preferred the mirrored mappings. A novice user wrote “[I am] used to mirror image interaction, because [I am] used to seeing objects move away when moving object/cursor closer to me.” Others, noted caveats, and cases where they thought the mirrored interaction would be useful. “In [the] second and third task [ISO tasks] iteration techniques C and D [mirrored] are also easy to use [larger object near the screen]. However A and B feels more natural.” (This person in fact did better in the mirrored cases for the static ISO task). “The mirrored tasks were more difficult but when the input wasn’t mirrored my interactions occluded the display.”

In terms of the skew differences between the mappings most participants had less to say: “To be honest I couldn’t really tell the difference from the previous method” (referring to the two non-mirrored cases). However, some participants—in particular participants with a lot of 3D gaming experience—preferred perspective mapping: “I preferred to have the cursor move in relation to the box” (i. e., perspective; a very experienced gamer). Others, usually novices, stated a preference for orthographic mapping: “intuitive, felt very much like using a mouse,” “felt the most natural.”

6 DISCUSSION

The results of the study show quite a large variation in response between participants, indicating that all examined mappings are in fact viable and possible alternatives when compared to the de facto standard of perspective mapping. They also show how adaptable humans are to change in control-display mappings and conflict between the visual and sensory-motor models, as previous work suggests [6]. The remarks given by our participants and the ranking results suggest that participants have clear personal mapping preferences and that the standard mapping is not necessarily their preferred first choice. The mean completion times show a significant difference between mirrored and non-mirrored mappings for all tasks. However, for all tasks the variation in mean times is high relative to the actual difference in means, indicating a small practical significance of the effect. From our trace analysis we saw that participants had difficulties in all mappings, even in the mappings in which they performed best. Most of these problems occurred in the depth dimension (z) as was to be expected from previous study results [10, 26] but were generally quite accurate in x and y , especially in the ISO tasks. When they had difficulty in z they attempted to adjust, likely causing the observed increased motion and difficulty in height (y) and from time to time in width (x) as well.

Although we used the perspective box with a grid pattern to provide a strong perspective visual cue, in all cases the performance in the orthographic mapping was not significantly worse than the perspective mapping. This and the fact that the orthographic mapping was the favourite mapping of the majority of participants is surprising since most 3D rendering uses perspective projection. It remains to be seen whether a lack of such strong perspective cues or the addition of other depth cues would support the de facto standard perspective mapping or whether orthographic mapping and other skew variations might be more appropriate. The mirrored interactions are new and had almost no visual support at all in the rendering system, so they may be much more viable than they appear now if appropriate support and visual cues were added.

The differences between results in the different tasks suggest that either certain mappings may be more appropriate for certain tasks or that there is a learning effect. In the first task (grid) participants did best in the non-mirrored orthographic mapping. When reaching the last analyzed task (moving ISO) the variation was reduced to a minimum. This could suggest that there is little learning needed to use non-mirrored orthographic mapping. In non-mirrored perspective mapping the grid task was completed slightly slower, in particular for novice users. Even with the visual cues supporting it, participants do approximately as well as when using the non-mirrored orthographic mapping in the moving ISO task. Perhaps the most surprising result is that performance in the mirrored mappings, which have almost no visual support, improves from the grid task to the moving ISO task and that this improvement is larger than in the non-mirrored mappings. If visual cues to support the mirrored mappings were introduced, it may be that these mappings are even more viable than they appear in this study.

7 CONCLUSION

We provide a general and mathematical control-display description with our conceptual framework. It includes mappings for scale, order, lag, orientation, flip, and skew. To explore the effects of skew and flip we conducted a study on four mappings: orthographic, mirrored orthographic, mirrored perspective and the de facto standard perspective. Our study results indicate that all three non-standard mappings may be considered as viable alternatives. Surprisingly, the de facto perspective mapping was not the favourite mapping for the majority of participants. Even more surprising is that, with little visual support, performance in the orthographic mapping was not significantly worse, and that performance in the mirrored mappings steadily improved even though there is almost no visual support for these mappings. In summary, our conceptual framework and the results from our study of four control-display space mappings together open the door to a new range of exploration into 3D interaction space mapping variations.

ACKNOWLEDGEMENTS

We would like to thank Natural Science and Engineering Research Council of Canada, Alberta's Informatics Circle of Research Excellence, Alberta Ingenuity, and the Canadian Foundation of Innovation for research support. We also thank our fellow researchers from the Interactions Lab at the University of Calgary for their insightful comments on this work.

REFERENCES

- [1] M. Agrawala, D. Zorin, and T. Munzner. Artistic Multiprojection Rendering. In *Proc. of EGWR 2000*, pages 125–136, London, UK, 2000. Springer-Verlag.
- [2] R. Balakrishnan. Why Aren't We Using 3D User Interfaces, and Will We Ever? In *Proc. of 3DUI 2006*, page 141, Los Alamitos, CA, Mar. 2006. IEEE Computer Society.
- [3] R. Blanch, Y. Guiard, and M. Beaudouin-Lafon. Semantic Pointing: Improving Target Acquisition with Control-Display Ratio Adaptation. In *Proc. of CHI 2004*, pages 519–526, New York, 2004. ACM Press.
- [4] J. Boritz. *The Effectiveness of Three-Dimensional Interaction*. PhD thesis, University of British Columbia, Canada, 1999.
- [5] J. Boritz and K. S. Booth. A Study of Interactive 6 DOF Docking in a Computerized Virtual Environment. In *Proc. of VRAIS 1998*, pages 139–146, Los Alamitos, CA, 1998. IEEE Computer Society.
- [6] E. Burns, S. Razzaque, A. T. Panter, M. C. Whitton, M. R. McCallus, and J. Frederick P. Brooks. The Hand is Slower than the Eye: A Quantitative Exploration of Visual Dominance Over Proprioception. In *Proc of VR 2005*, pages 3–10. IEEE, 2005.
- [7] L. Carroll. *Through the Looking Glass and What Alice Found There*. MacMillan and Co., Ltd., London, UK, 1872.
- [8] P. Coleman and K. Singh. RYAN: Rendering Your Animation Nonlinearly projected. In *Proc. of NPAR 2004*, pages 129–138, New York, 2004. ACM Press.
- [9] S. Frees and G. D. Kessler. Precise and Rapid Interaction Through Scaled Manipulation in Immersive Virtual Environments. In *Proc. of VR 2005*, pages 99–106, 2005.
- [10] T. Grossman and R. Balakrishnan. Pointing at Trivariate Targets in 3D Environments. In *Proc. of CHI 2004*, pages 447–454, New York, 2004. ACM Press.
- [11] T. Grossman and R. Balakrishnan. An Evaluation of Depth Perception on Volumetric Displays. In *Proc. of AVI 2006*, pages 193–200, New York, 2006. ACM Press.
- [12] I. S. MacKenzie and S. Riddersma. Effects of Output Display and Control-Display Gain on Human Performance in Interactive Systems. *Behaviour & Information Technology*, 13(5):328–337, 1994.
- [13] A. D. Milner and M. A. Goodale. *The Visual Brain in Action*, volume 2 of *Oxford Psychology Series*. Oxford University Press, New York, 1995.
- [14] B. A. Po, B. D. Fisher, and K. S. Booth. Pointing and Visual Feedback for Spatial Interaction in Large-Screen Display environments. In *Proc. of Smart Graphics 2003*, volume 2733 of *Lecture Notes in Computer Science*, pages 22–38, Berlin, Heidelberg, 2003. Springer-Verlag.
- [15] E. C. Poulton. *Tracking Skill and Manual Control*. Academic Press, New York, 1974.
- [16] I. Poupyrev and E. Kruiff. 20th Century 3DUI Bib: Annotated Bibliography of 3D User Interfaces of the 20th Century. 3rd revision, 2000.
- [17] I. Poupyrev, S. Weghorst, and S. Fels. Non-Isomorphic 3D Rotational Techniques. In *Proc. of CHI 2000*, pages 540–547, New York, NY, USA, 2000. ACM Press.
- [18] B. Shneiderman. Why Not Make Interfaces Better than 3D Reality? *Computer Graphics and Applications*, 23(6):12–15, Nov./Dec. 2003.
- [19] R. W. Soukoreff and I. S. MacKenzie. Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts' Law Research in HCI. *International Journal of Human-Computer Studies*, 61(6):751–789, Dec. 2004.
- [20] R. Stoakley, M. J. Conway, and R. Pausch. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proc. of CHI 1995*, pages 265–272, New York, 1995. ACM Press/Addison-Wesley Publishing Co.
- [21] C. Swindells, B. A. Po, I. Hajshirmohammadi, B. Corrie, J. C. Dill, B. D. Fisher, and K. S. Booth. Comparing CAVE, Wall, and Desktop Displays for Navigation and Wayfinding in Complex 3D Models. In *Proc. of CGI 2004*, pages 420–427, Los Alamitos, CA, 2004. IEEE Computer Society.
- [22] Y. Wang and C. L. MacKenzie. Object Manipulation in Virtual Environments: Relative Size Matters. In *Proc. of CHI 1999*, pages 48–55, New York, 1999. ACM Press.
- [23] L. C. Wanger, J. A. Ferwerda, and D. P. Greenberg. Perceiving Spatial Relationships in Computer-Generated Images. *IEEE Computer Graphics and Applications*, 12(3):44–58, May 1992.
- [24] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Effects of Display Position and Control Space Orientation on User Preference and Performance. In *Proc. of CHI 2006*, pages 309–318, New York, 2006. ACM Press.
- [25] S. Zhai. User Performance in Relation to 3D Input Device Design. *SIGGRAPH Computer Graphics*, 32(4):50–54, Nov. 1998.
- [26] S. Zhai, P. Milgram, and A. Rastogi. Anisotropic Human Performance in Six Degree-of-Freedom Tracking: An Evaluation of Three-Dimensional Display and Control Interfaces. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 27(4):518–528, July 1997.