

# Towards an Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration

David López, Lora Oehlberg, Candemir Doger, and Tobias Isenberg, *Senior Member, IEEE*

**Abstract**—We discuss touch-based navigation of 3D visualizations in a combined monoscopic and stereoscopic viewing environment. We identify a set of interaction modes, and a workflow that helps users transition between these modes to improve their interaction experience. In our discussion we analyze, in particular, the control-display space mapping between the different reference frames of the stereoscopic and monoscopic displays. We show how this mapping supports interactive data exploration, but may also lead to conflicts between the stereoscopic and monoscopic views due to users' movement in space; we resolve these problems through synchronization. To support our discussion, we present results from an exploratory observational evaluation with domain experts in fluid mechanics and structural biology. These experts explored domain-specific datasets using variations of a system that embodies the interaction modes and workflows; we report on their interactions and qualitative feedback on the system and its workflow.

**Index Terms**—Visualization of 3D data, human-computer interaction, expert interaction, direct-touch input, mobile displays, stereoscopic environments, VR, AR, conceptual model of interaction, interaction reference frame mapping, observational study.

## 1 INTRODUCTION

EFFECTIVE visualization and interaction with 3D data that represents scans or simulations of the real world is at the heart of visualization research. Exploration and analysis are most strongly supported when combining the best possible visual representations with the best possible interaction techniques. Previously, researchers have primarily focused on improving visual representations; more recently visualization research has increasingly emphasised suitable interaction techniques. Most approaches present visual representations to users by targeting traditional PC-based settings, while some also explore virtual environments [14]—in particular, the use of stereoscopy facilitates depth perception and thus ***high visual immersion***. To enable interactive exploration of visualizations, most approaches again target PC-based settings while some use novel devices such as horizontal/vertical large displays or mobile devices. In particular, the use of touch-based interaction for visualization provides ***high immersion through interaction*** due to its directness—the touch input and the affected data are at the same visual location (i.e., sticky interaction [28]), resulting in users feeling “in control of the data” [75, 79].

Unfortunately, these two ways of achieving immersion are mutually exclusive. On the one hand, virtual objects in stereoscopic settings cannot be touched since they appear to float in empty space. Touch interaction, on the other hand, conflicts with stereoscopic dis-

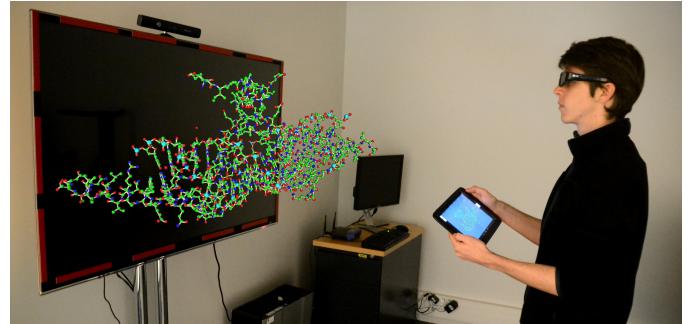


Fig. 1. Tablet-based navigation of stereoscopically displayed 3D data.

play due to parallax issues [13, 71, 72] as well as touch-through [16, 70] and invisible wall problems—it is far better suited for monoscopic displays. In visualization and HCI, researchers have thus explored hybrid settings that separate touch input from stereoscopic display. In addition to certain static setups targeted at interactive visualization (e.g., [9, 17, 18, 56]), researchers have investigated ways to control 3D scenes using orientation sensors and touch capabilities of mobile devices. These systems, however, often do not take full advantage of the sensation of direct manipulation offered by touch-enabled monoscopic displays or their ability to provide versatile, data-specific interface designs.

To enable researchers and engineers to build powerful hybrid visualization environments in the future, we therefore work towards a more nuanced understanding of the use of stereoscopic displays in this context—stereoscopic visualizations of 3D data that are controlled via touch interaction on a monoscopic view of the data on a mobile tablet. Our goals are to (a) allow both visual immersion (stereoscopy) on the one hand and immersion through interaction (direct touch) on the other hand; to (b) support the mobile viewing of and interaction with the data; and to (c) offer rich interaction through a lightweight, minimalist implementation.

To accomplish these goals we contribute an in-depth analysis

- David López is with Inria, France, and University of Antioquia, Colombia. E-mail: david.lopezb@udea.edu.co.
- Lora Oehlberg is with the University of Calgary, Canada. E-mail: lora.oehlberg@ucalgary.ca.
- Candemir Doger is with Inria, France, and Sabancı University, Istanbul, Turkey. E-mail: cdoger@sabanciuniv.edu.
- Tobias Isenberg is with Inria, France. E-mail: tobias.isenberg@inria.fr.

Manuscript received 3 Dec. 2014; revised 23 Mar. 2015; accepted 12 May 2015. Date of publication 0 . 0000; date of current version 0 . 0000. Recommended for acceptance by D.A. Bowman.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TVCG.2015.2440233

of the interaction workflow that arises when using tablet-based navigation of 3D datasets in a stereoscopic viewing environment (e.g., Fig. 1). In our analysis we revisit several interaction modes and propose ways for the user to fluidly move between these modes during exploratory visualization. To better understand this conceptual design, we implemented a system that facilitates seamless mode shifts as users move around and interact with a stereoscopic view via a monoscopic touch tablet. To demonstrate breadth, we created two variations of the system using two established touch-based tablet interaction techniques for visualization. We then conducted an exploratory qualitative evaluation of the viability of the interaction workflow by observing domain experts in fluid mechanics and structural biology as they explored and navigated through 3D datasets, teaching us not only about the general interaction design but also about the benefits and challenges of the specific navigation techniques in this 3D visualization context.

Our most important contribution, in summary, is a conceptual model of interaction modes that describe the task-dependent relationships between the visualizations on the stereoscopic display and the touch-based interactions to navigate and control the visualization on the monoscopic tablet. In particular, we describe the mapping between the different reference frames of the tablet and the stereoscopic data view as well as discuss in depth the implications of these not being synchronized. We explain why interaction is still easy in some situations despite reference frame differences, and we show when synchronization is necessary.

## 2 RELATED WORK

Interaction with 3D objects and data is an active research field with a long tradition [11, 29, 37, 41]. Much of this work is based on stereoscopy; researchers have created and studied virtual reality environments [14] (e.g., CAVE [20] and Responsive Workbench [45]). Interaction in VR settings often relies on 3D tracking and uses wands, gloves, haptic feedback, or special devices. While such input facilitates a direct 3D placement of objects (in addition to indirect input using rate control for larger-scale navigation), this general interaction approach suffers from the *touching the void* issue [12] (except for haptic feedback). This problem can only be partially addressed with transparent props [63], at the expense of the ability to move the touched objects.

Researchers have thus investigated other ways of adding somesthetic feedback [61] to the interaction with 3D scenes or data [32, 34], with or without stereoscopic depiction. The straight-forward way of combining stereoscopy with touch interaction is to capture touch input on the stereoscopic display surface. However, this method has numerous challenges including parallax, touching through empty space, and bumping into an invisible surface [13, 16, 33, 70, 71, 72]. Some approaches deal with these issues by adjusting the 3D world such that the surface of the target object moves toward the display surface when a touch occurs [69, 70]. While this technique works well when manipulating individual 3D objects, in visualization it is important to not distort a dataset's spatial arrangement to avoid incorrect perception of the data.

Another approach locates the touch input on a dedicated surface, separated from the stereoscopic projection. Toucheo [27] is an innovative system based on this principle. Its touch-sensitive surface shows the interface elements, while a semi-transparent mirror—placed *over* the touch surface with enough space for the interacting arms—adds the stereoscopic elements. People thus perceive their hands to be below the floating objects, while visual elements

connect the monoscopic and stereoscopic parts. However, the Toucheo setup limits people's freedom to move around and is thus only suitable for desk-like display settings.

Other researchers investigated the use of mobile monoscopic touch-enabled devices, such as phones or tablets, in a VR (or monoscopically projected 3D) context [57]—similar to the setup in our work. The interaction in these setups typically maps the mobile device's orientation sensors to the orientation of either the manipulated objects [22, 39, 51] or the scene/the camera [7, 22, 25]. Translations of data/the camera in these settings are usually achieved by touch-dragging on the mobile device (e.g., [42]), via linearly scaled direct offsets, flicking, or in form of rate control. Similarly, zooming is often done using a pinching gesture on the touch surface. Others simulate buttons/joysticks [24, 50] or use the mobile device as a touch-sensitive handle/widget that is virtually extended into the scene for manipulations [44, 48]. One special form of interaction is the use of a touch-sensitive mobile device whose surface is mapped to an arbitrarily shaped virtual surface in 3D [58] such that touch interactions can be used to manipulate 3D objects. All of these (3D navigation) techniques<sup>1</sup> share the disadvantage that they do not take the mobile device's or its user's position into account, thus restricting the viewer position to a single front-on view—we address this issue in our work.

A widget-based approach for manipulating 3D objects at a distance is SquareGrids [51] which uses dedicated regions for separating on-object, off-object, and environment interactions. The general use of widget-based interaction over purely gestural patterns is beneficial in visualization applications [35] because it facilitates spring-loaded interaction design [65]. The many interaction techniques needed in 3D data exploration can, therefore, be integrated into a usable interface that is easily remembered by people using it. A good example of this principle is the Slice WIM setup [18], the setup most related to our own as it uses both touch input and stereoscopic viewing for interactive visualization. Slice WIM combines a vertical stereoscopic screen for a detailed data view with a statically attached, monoscopic horizontal display that captures touch input and displays interaction widgets. Using this unique setup, Coffey et al. [18] display a stereoscopic world-in-miniature (WIM) of the dataset being explored with representations of the currently controlled detail view such that scientists can, for example, position cutting surfaces in 3D. While its separation of touch surface from stereoscopic display surface nicely avoids stereoscopic touch issues, the Slice WIM setup is static and restricts people's position with respect to the stereoscopic display. In our work we remove this restriction, explore the resulting interaction mapping issues, and propose solutions for them.

## 3 TWO-DISPLAY DATA NAVIGATION

Our overall vision is to enable researchers to explore 3D datasets with as much *immersion* as possible, arising both *from visuals* as well as *from interaction*. We therefore explore ways to combine a stereoscopically displayed large view of the 3D data with direct touch input to control the view. This combination has the potential to increase people's acceptance of stereoscopic environments for 3D data visualization since—through touch-based interaction—it puts them in control of their data (e.g., [79]). Moreover, the indirect manipulation of (stereoscopically displayed) 3D data on a personal

<sup>1</sup> There is, in fact, some work on spatially-aware 2D navigation using a mobile device in front of a large display (e.g., [30, 59]). In our work, however, we focus on 3D interaction techniques, specifically 3D navigation.

touch device was shown to have potentially more efficient and precise interaction than direct interaction on a large display [44]. In addition, as pointed out in Sect. 2, capturing touch input on the stereoscopic display’s surface introduces several issues that prohibit intuitive interaction. We thus investigate settings that separate touch input from the stereoscopic display by using a mobile device (a tablet) for both touch input and additional visual feedback.

### 3.1 Narrowing Down the Design Space

The design space for achieving high visual immersion through stereoscopically displayed data and high immersion from interaction on mobile touch devices is quite large. We thus begin to define this space through a series of design guidelines, motivated by our previous experience with creating interactive visualization systems.

To begin, the fundamental need for at least 7DOF navigation is well established for 3D data visualization:

G1: The system should support 7DOF navigation of the dataset: 3D translation, 3D rotation, and uniform scaling.

These 7DOF should be controlled consistently to avoid breaks in the interaction scheme. Moreover, consistency is crucial for integrating navigation with additional data exploration tools to ultimately provide a complex visualization environment [40]:

G2: Navigation interactions should allow for consistent combinations of (potentially many) data exploration techniques.

The third design guideline arises from the need to associate navigation input events with specific parts of the 3D data, as well as from the need for flexible interaction design:

G3: The mobile device should actively display content. Both data and interaction elements should not only be shown in the large stereoscopic view but also displayed and actively used on the monoscopic mobile device that provides the interaction input.

Alternatively, the mobile device could be used as a touch-pad in a eyes-free manner. This design, however, would still have significant disadvantages. Without comprehensive tactile feedback (far beyond what is possible with today’s technology [38, 46, 67]) we would (1) be restricted to only indirect interactions, increasing the complexity of the interaction design (i.e., requiring complex gesture languages that are difficult to remember). The types of interactions that can be initiated on the mobile device are also still limited, as (2) data-specific input or widget-based designs are not possible. Finally, we would (3) need to integrate additional visual elements such as 3D cursors in the stereoscopic screen (i.e., 3D cursors) when selecting data. These additional elements both add complexity to the view design and increase rendering times since additional occlusions would need to be calculated [62].

In contrast, by displaying interaction widgets and data on the monoscopic mobile device, navigation interactions are performed as actions on the data itself, bringing greater ***immersion through interaction***. Through active widgets this design approach also facilitates a much ***richer set of interaction input*** and lets us make use of established direct manipulation techniques on monoscopic views for 3D navigation (e.g., [19, 79]) as well as, in the future, for more complex interactions such as selection and picking (e.g., [52, 77, 78]). This guideline thus favors larger, tablet-sized mobile devices due to their larger display size that can visually display more of a dataset compared to, for example, a smartphone.

To fully benefit from a stereoscopic viewing environment, the viewer should be able to intuitively observe and analyze the data

from different perspectives by physically changing their position with respect to the stereoscopic data display:

G4: The interacting person should be able to move around the space in front of the stereoscopic display.

This fourth guideline, paired with stereoscopic viewing, implies 3D tracking the position of the interacting person. While there are many ways to implement tracking, simplicity and robustness are essential for the practical success of any visualization system:

G5: 3D tracking should be as simple as possible.

Simple tracking could mean, e.g., the use of less-elaborate tracking systems (e.g., time-of-flight 3D cameras) that do not require people to wear or carry extra tracking equipment—in particular on the mobile device. Such simplified tracking environments that stand *today*’s practical test are also more easily accepted by our target audience of domain experts due to their simplicity, versatility, and low maintenance needs (e.g., [31, 53, 68]).

Unique in scientific data exploration is the need to precisely control the data analysis process to obtain accurate results [31]:

G6: Input should be well-controlled and as precise as possible.

Precise control of visualization applications is challenging, in particular with touch input [43]. Issues arise not only from touch input (content occlusion and the fat finger problem [6]) but also from the need to separate the different degrees of freedom [55].

Finally, we re-state a goal common to most touch-based interaction on mobile devices to reduce excess fatigue:

G7: Fatigue issues arising from physically holding a touch-interactive mobile device should be kept to a minimum.

This implies that people should be able to adopt a range of positions with the mobile device as they interact with it. Requiring people to be constrained to awkward positions (e.g., continuously hold the mobile device in mid-air) would quickly tire their arms.

Together, these guidelines imply that we focus our investigation on stereoscopic displays paired with touch-interactive tablet-sized mobile displays. A tablet is large enough to show interaction widgets as well as a reasonably-sized version of the visualization on the display (G3) and, therefore, provide more precise input than on a phone (G6), while still being lightweight enough to be held during interaction (G4, G7). In our stereoscopic environment the person is tracked, but not the tablet (G5).

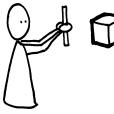
### 3.2 Interaction Modes for Tablet-Based 3D Navigation

Starting from these design guidelines, we now discuss possible interaction modes for the touch-based mobile device that support navigation of 3D data visualizations. We support this discussion of a possible interaction model with Table 1, explaining the limitations of its ‘intuitive’ interaction approach. We begin with guideline G3: to ultimately enable rich data exploration we want to show visualizations both stereoscopically on the large display and monoscopically on the tablet. One common way that people experience both monoscopic and stereoscopic views is when using digital cameras—or tablets as digital cameras—in the physical world: as they take pictures of objects, people see both a stereoscopic view of a (real-world) object as well as the monoscopic projection produced by the camera. This setup also resembles how many interaction techniques in augmented reality (AR; e.g., [3, 15]) are designed to view a virtually enriched view of the natural world.

Compared to both these environments (physical world and AR), the *exploration of 3D data* displayed on a stereoscopic screen is different in that *all* visualized data is virtual—the physical world

TABLE 1

Initial tablet interaction design considerations, explaining the limitations of a straight-forward AR-based approach to a hybrid tablet+stereo display interaction setting. For each interaction mode, its purpose, the viewer's and tablet's virtual cameras as well as the physical tablet positions are compared, and the resulting ease of mapping/handling are given. The arrows indicate the typical initial AR-inspired workflow.

tablet operation metaphor/mode	A: video camera (taking pictures)	B: still camera, after picture	C: still camera, touch interaction	D: still camera, gyroscope	E: still camera, viewer moves
<b>purpose of the interaction</b>	set/re-set correct tablet view	obtain controllable & stable interaction setting	clear & controlled data navigation	small data orientation adjustments	view adjustments in 3D data space
<b>viewer camera</b>	tracked; = viewer	tracked; = viewer	tracked; = viewer	tracked; = viewer	tracked; = viewer
<b>tablet camera</b>	≈ viewer camera; ≈ viewer (inferred)	≈ viewer camera; ≈ viewer	≈ viewer camera; ≈ viewer	≈ viewer camera; ≈ viewer	≠ viewer camera; ≠ viewer
<b>tablet position &amp; orientation w.r.t. cameras</b>	= tablet camera; ≈ viewer camera	lower than tablet/viewer cameras, x-rotated ≈ 60°	lower than tablet/viewer cameras, x-rotated ≈ 60°	lower than tablet/viewer cameras, <b>1–3 rotations</b>	lower than tablet/viewer cameras, <b>2 rotations</b>
<b>mapping clarity</b>	clear	clear	clear	<b>unclear</b>	<b>unclear</b>
<b>interaction focus</b>	tablet	tablet & stereo view	tablet & stereo view	stereo view	tablet & stereo view
<b>tablet handling</b>	awkward	easy	easy	moderately easy	easy
<b>initial workflow</b>					

only serves as a reference frame for the (virtual) location of the data. Nevertheless, we can use the metaphor of the tablet as a physical object that controls a *virtual (video) camera* in the virtual data space, looking at the 3D dataset (Table 1, column A). In this case, view direction and position are tracked through built-in orientation sensors and the virtual camera's absolute position is inferred by tracking the person holding it (G5). This interactive setup specifically resembles handheld augmented reality interfaces (e.g., [60, 64, 73]) and also relates to camera-phone-based interactions in multi-display environments (e.g., [10]).

However, the video camera metaphor has fundamental usability problems. First, as the person needs to direct the camera toward the target objects, the tablet must be held at head height with out-stretched arms to obtain and retain a good view of the dataset. The person's arms will easily get tired, thus contradicting G7.

Second, and even more importantly, the video camera metaphor implies a continuously updated view on the mobile device which causes serious issues with respect to the fatigue (G7) and precision/control (G6) guidelines: An interaction mapping with constant updates would, for any meaningful and controlled navigation input, require the user to hold and maintain the tablet at a fixed orientation and position for the duration of the interaction. This fixed position is essential to maintain a steady frame of reference to issue precise navigation commands, yet the inherent inaccuracy of sensor-based 3D tracking and the inability of humans to remain completely still prevent such exact input.

While in see-through AR the continuous updates arise from an actual camera showing a constantly changing image of the physical world, our completely virtual data display is not bound by such constraints. We can reduce fatigue and obtain better interaction precision by no longer tracking the tablet's view of the virtual data. Instead, we propose to use the metaphor of a *digital still image camera* in which *the view on the tablet is only updated on-demand*. Essentially, after the camera has a good view of the data, the person takes a picture and holds the tablet at a more comfortable, lowered position to examine the image (Table 1, column B). A similar interaction paradigm was also suggested in handheld augmented reality [26, 49] (e.g., “Freeze-Set-Go”) and multi-display [10]

settings (“freezing”). In both our visualization setting and in the AR/multi-display scenarios, freezing and the subsequent relaxation of the tablet's position serves both G6 and G7: it provides a *stable reference frame for touch interaction* and makes the handling of the tablet much more comfortable and less fatiguing.

However, there is a key difference between AR/multi-display environments and our setting with a 3D visualization on a stereoscopic display. For handheld AR/multiple displays, the view on the tablet includes both a static representation of the physical world (an image taken by a camera on a mobile device) and a virtual overlay. In our case, in contrast, both the stereoscopically displayed 3D scene and the tablet's view are entirely virtual. This difference means that, unlike in handheld AR, we can still see and control the entire 3D scene (the visualization)—the person examining a dataset can *fully control the view of the visual representation*, either through touch input (Table 1, column C) or gyroscopic interaction (Table 1, column D) with the tablet. It is essential to note that the views on the tablet and the stereoscopic display remain fully linked: both are being updated and both can be used for visual feedback for 3D interaction. Moreover, it is important that this interactive control can and needs to happen for all 7DOF of the 3D visualization (G1), in contrast to the frozen 2D interaction with documents in multi-display environments [10].

### 3.3 Mobile to Stereoscopic Reference Frame Mapping

Because the tablet's projected view and the stereoscopic view are continually linked, we must not only carefully consider the touch-based 3D navigation interface, but also understand the specific *mapping between the reference frames of the tablet and the stereoscopic 3D view*.

Let us first consider the case of a static viewer who uses the tablet at a relaxed, frozen position (Table 1, columns B–D). In this case, we may want to employ established touch-based 3D navigation techniques such as the tBox [19] or FI3D [79] widgets on the stable, monoscopically displayed view on the tablet. Then, as we affect the 3D position, orientation, and scale of the dataset, both virtual cameras—the viewer's and the tablet's—can remain

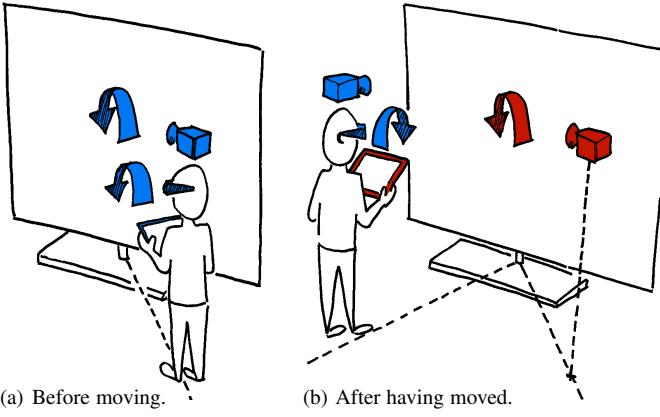


Fig. 2. Virtual cameras & physical tablet locations. Before moving (a), the views are synchronized (blue); after moving (b), two separate views/reference frames (blue and red) exist and interactions on the tablet lead to views of the data which are confusing to the interacting person.

at their respective, identical or roughly identical locations in the 3D visualization space (Fig. 2(a)). However, the (touch) interaction is still clear<sup>2,3</sup> to the person navigating the 3D visualization space (row ‘mapping clarity’ in Table 1), despite the physical location of the tablet being offset from the virtual position of the tablet camera (row ‘tablet position & orientation with respect to cameras’ in Table 1). Instead of being identical to its virtual camera, the physical tablet is lower and is typically rotated around the  $x$ -axis—oriented somewhat like a drafting table.

This mapping from tablet touch input to the resulting manipulations of the stereoscopically observed dataset thus resembles people’s way of controlling desktop environments using mice—studied within the context of sensorimotor adaptation [76]: 2D mouse motions on the horizontal control space are mapped to 2D motions of the cursor on the vertical display space. For a vertical display, this ‘mouse mapping’ comprises a 90° rotation in pitch plus a horizontal and a vertical offset [80] and is familiar to most people who frequently interact with PCs.

What is particularly important from this sensorimotor adaptation research in our context of 3D interaction with visualizations is that this ‘mouse mapping’ from control space to display space works well **only** as long as the involved transformations are limited to roughly the described  $x$ -rotation, potentially with an additional translation. This is the case because translations are easy to adapt to [4, 8, 76]; and rotations are somewhat more difficult [21] but still reasonably easy **for a single rotation** with an angle of up to 120° [1, 2]. In the mentioned ‘mouse mapping’ case, the  $x$ - and  $y$ -axes were shown to depend on one another [5]. With an additional rotation around the  $y$ -/ $z$ -axes, however, the frames of reference can no longer be easily aligned through a mental rotation and thus control becomes more difficult (Fig. 2(b)). In fact, a similar dependency of frame of reference mappings for rotations around the  $y$ -axis was described for the control of 3D objects in stereoscopic display environments, with increasing interaction difficulty when the angle of rotation increases [74].

This is why, when the *user moves around the stereoscopic display* in addition to holding the tablet at the frozen, relaxed

2. At this point we make this claim of clarity for the interaction mapping based on observations by us as authors; later in the paper we report on an empirical evaluation of this mapping with experts.

3. During a gyroscope interaction, the mapping is also not clear due to the several involved rotations (see Table 1), but there this is not a problem because the focus of the viewer is on the stereoscopic display and not on the tablet.

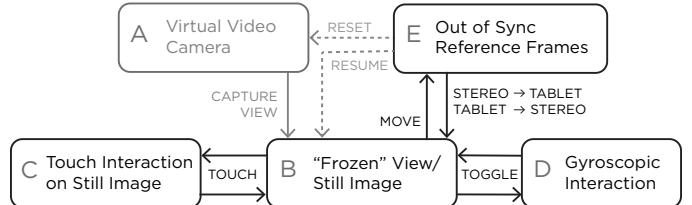


Fig. 3. Final workflow of tablet-based navigation of stereoscopic visualizations. The gray and dashed elements are not recommended.

(rotated and offset) position, the workflow no longer works because what is shown on the tablet (when interacting) conflicts with what is observed from the stereoscopic display: the data view users see through their eyes is adjusted through 3D tracking, yet the tablet’s view is static due to the still camera metaphor (Table 1, column E). In our setting we can thus expect a person exploring a visualization to have increasing difficulties with navigating a 3D dataset once he or she starts moving away from the point at which the tablet view and the stereoscopic view were synchronized. A rotation around the  $x$ -axis, for instance, leads to a simple rolling action on the tablet’s view of the dataset but is shown as rotation around a different axis in the stereoscopic view observed on the large display.

### 3.4 Reference Frame Synchronization

The important insight of these considerations on the reference frame mapping is that, at some point during interactive data exploration, the arising discrepancy needs to be resolved and the two views need to be re-synchronized. One way to accomplish this synchronization is to return back to the virtual video camera interaction mode after a move to a new position. Resolving the mismatch between tablet view and stereoscopic view by forcing the user to return to the video camera metaphor, however, is awkward; it discards any previous useful reference frame and requires a complete mental restart of visual analysis. In a VR setting such as ours, however, the tablet’s video camera is not limited by physical reality to get a good view of the 3D dataset—we already know the best possible camera position from 3D tracking. We can thus directly use this tracking data to re-sync the tablet’s view.

At this point one may think that an easy solution would be to go back to a constant synchronization between tablet and stereoscopic view. As argued in Sect. 3.2, however, such a design would cause the tablet view to constantly change, contradicting guideline G6 that asks for control and precision. Moreover, we would not be able to capture a good view using the tablet and easily go back to it or show it to collaborators who may be viewing the same dataset from different positions around the stereoscopic display. Therefore, we use two dedicated and straight-forward synchronization mechanisms: *stereo view → tablet* and *tablet → stereo view*. In the former (*stereo view → tablet*), the stereoscopic camera parameters from 3D tracking are used to define the tablet’s virtual camera. As a result, the tablet shows the exact same view that the user sees stereoscopically at the time of the interaction. The latter (*tablet → stereo view*), supports the opposite effect: the stereo view is adjusted such that the interacting person sees the same VR view as was previously displayed only on the tablet. Since we cannot change the virtual camera that computes the stereoscopic view, we re-arrange the dataset such that the viewer camera shows the same view as previously shown only on the tablet.

The use of the dedicated synchronization mechanisms means that the relatively awkward physical action from the video camera

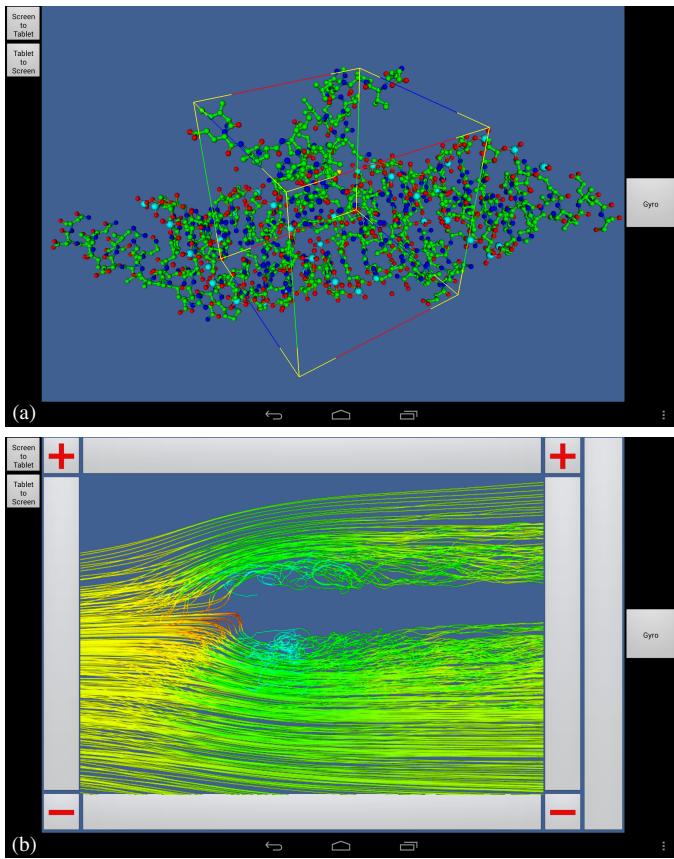


Fig. 4. Tablet interface: (a) tBox and (b) FI3D.

metaphor is replaced by the two user-issued commands—with the additional advantage that with both commands the tablet view exactly matches the stereoscopic view, as opposed to manually approximating the same view via the video camera metaphor. The resulting workflow (Fig. 3) satisfies all our design guidelines.

### 3.5 Prototypical Implementation and Setup

We built a prototype system that embodies this workflow to evaluate our interaction design. We use the VTK library<sup>4</sup> to render the stereoscopic view for the large display (with off-axis rendering [23]), and the VES library<sup>5</sup> to display the dataset on the tablet. Both devices are linked using a client-server setup in which the tablet acts as the client and the stereoscopic display application is the server. When the tablet detects input, it sends the data transformation to the server via wifi. As shown in Fig. 1, the stereoscopic view is shown on a 3D TV with a  $1920 \times 1080$  pixel resolution (55 inch = 139.7 cm diagonal) using active shutter glasses. The simple tracking (see G5) is achieved with a Microsoft Kinect and the additional manual setting of an interacting person's eye height. The tablet we use is a Google Nexus 10 (pixel resolution:  $2560 \times 1600$ , display diagonal: 10.1 inch = 25.5 cm, spatial device size: 263.9 mm  $\times$  177.6 mm  $\times$  8.9 mm, weight: 603 g).<sup>6</sup>

For navigation, we implemented two different techniques to explore a range of touch-based interaction. We decided to use tBox [19] and FI3D [79] (see Fig. 4) over other touch-based interaction

techniques (for an overview see Isenberg's work [34]) because they both delivered at least 7 DOF control and could be used with finger input and without an additional stylus. Both tBox and FI3D were implemented as described in their original papers, with slight adjustments. To be useful for data space navigation (rather than manipulating individual objects in a 3D scene), we maintain the tBox centered at the origin of the virtual scene and only allow uniform two-finger scaling instead of the non-uniform scaling possible with the original tBox. For FI3D we selected the one-finger interaction to map to arcball rotation [66] because for the kinds of datasets we experimented with it seemed that rotation was the primary interaction technique used in exploration. To conserve space, we removed FI3D's top and bottom bars for translation along the z-axis and instead placed a vertical bar for this purpose to the right of the regular frame. For both the tBox and the FI3D widgets, we left space on the left and the right for the view synchronization buttons and the spring-loaded gyroscope interaction. We implemented the gyroscope control by using Android's virtual rotation sensor that is based on the tablet's gyroscope, magnetometer, and acceleration sensors, combined in a Kalman filter. The resulting tBox, FI3D, and gyroscope manipulations are applied directly to the dataset's model-view transformation.

## 4 OBSERVATIONAL STUDY WITH EXPERT USERS

To better understand how our design would fare in practice we conducted an observational study with domain experts. In visualization, there are many different evaluation goals and types to choose from [36, 47]. In our case we were less interested in algorithmic performance, a qualitative results inspection, or user performance with respect to time and error metrics. Instead, we wanted to explore how domain experts deal with the inherent interaction mapping problems when using tablet-based navigation without synchronization, how our interaction concept would help them analyze 3D data, and what we can learn about the design of touch-based 3D data navigation in a stereoscopic environment. We were specifically *not* interested in whether the tBox or FI3D performed better in this very specific but typically artificial, controlled setting. Instead, we wanted to qualitatively compare the two interfaces, see what types of interaction goals they do or do not support well, and how we can create a better interaction designs for future interactive visualization environments. In addition, while both methods were originally designed as direct manipulation navigation techniques, we were interested in use in our indirect interaction scenario. After an intensive discussion on the most appropriate study methodology we, therefore, decided to conduct a *qualitative, observational study with domain experts*.

### 4.1 Participants

We recruited researchers from two labs whose work focuses on spatial 3D data: *structural biology* (analysis of biological molecules; 2 female, 2 male; M1–M4) and *fluid mechanics* (analysis of 3D flow data; 4 male; F1–F4). Participants' ages ranged from 24–60 (median: 31.5 years), with between one and 33 years of post-Master's level experience (median: 9.5 years) in the respective fields. All participants volunteered their time and were not compensated for their participation.

All of our participants interact with 3D data weekly or more frequently, typically with mouse-based interfaces. Space mice were used only by three participants, and then only infrequently, similar

4. See <http://www.vtk.org/>.

5. See <http://www.vtk.org/Wiki/VES>.

6. One may argue that observing tablets through shutter glasses is problematic due to the polarization in the glasses and in LC displays. This is not an issue, however, when using AMOLED-based devices such as the Google Nexus 10.

for force-feedback devices or 3D tracking ( $4 \times$ ). Stereoscopic environments had been used by five participants, although only by one weekly or more frequently. Touch interaction was more familiar to our participants with 7 out of 8 participants using it at least daily or more frequently. Most interact with touch interfaces on tablets on a monthly basis or more frequently. None of our participants reported interacting with large vertical (“wall”) or horizontal (“table”) displays, although three had participated in experiments and viewed data in a CAVE environment.

## 4.2 Apparatus

In addition to the setup described in Sect. 3.5, we used domain-specific datasets as well as a study controller.

**Domain-specific datasets.** We asked our participants ahead of time to provide us with datasets for the study, or to confirm datasets that we suggested. For each domain, we selected two datasets. For the fluid mechanics researchers, one dataset represented a snapshot of flow over a thick flat plate (Fig. 4(b)), where color maps to mass density. The other dataset displayed the atmospheric air motion over Europe, where color maps to the air temperature (Fig. 5). Both of these datasets showed streamlines of a single time frame. For the structural biologists, one dataset represented the tomato aspermy virus protein 2b (PDB: 2ZI0; Fig. 4(a)); the other dataset showed the E. coli WrbA holoprotein (PDB: 3ZHO). Both molecular datasets were shown as balls-and-sticks models, with the atoms colored according to the respective elements.

**Study Controller.** During the study, one experimenter operated the Study Controller. With this tool we were able to control the tablet interface, set the stereoscopic view to be tailored to the participant’s eye height and pupil distance, load datasets, change the participant’s stereoscopic viewpoint, and ‘mute’ the visual output of the stereoscopic display as people walked from one viewpoint to the next. The study controller also captured an action log as well as the participants’ interactions during each session.

## 4.3 Study Design and Tasks

We conducted structured observations of domain experts using different tablet interaction techniques to navigate domain-specific 3D datasets. At the beginning of each session (which each lasted between 60 and 90 minutes), our domain expert participants filled in an initial questionnaire to identify demographic information, their background in interacting with 3D datasets, and their familiarity with stereoscopic and touch-based interaction. We also measured their eye height and pupil distance to enter into the Study Controller and set up the stereoscopic view.

### 4.3.1 Training

Next, we introduced the interface. We instructed the participant on how to use the first interaction technique—either tBox or FI3D—to rotate, translate, and scale the data and introduced the gyroscope-based rotation. We encouraged participants to use each interaction until they felt comfortable and were able to reproduce specific views of the training data (the Stanford Bunny). After each completed view, the training data was reset to the same starting point.

### 4.3.2 First Section

After completing the interface training, the experimenter loaded one of the domain-specific datasets. We then asked the participant to stand directly in front of the stereoscopic display (point A, see Fig. 5; approx. 2 m from the display), and to explore the dataset

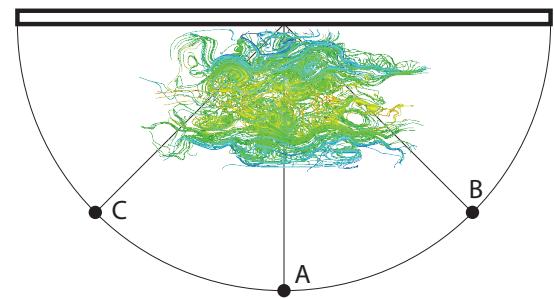


Fig. 5. General study setup with positions A–C, using the atmospheric air motion data used in the study as the example.

until they could identify a view that they found particularly interesting. We asked them to talk aloud, and to verbally indicate when they were done—this state was captured by the Study Controller.

We then asked the participant to move to an oblique point (point B). From this location, their stereoscopic view of the data changed to the viewpoint from point B; however, the tablet’s view was not modified and thus remained at point A. We asked the participant to reproduce their chosen view as they had seen it from point A. At this point, we expected many participants to experience difficulty due to the lack of synchronization. Afterwards, if the participants had not already voiced any difficulties on their own, we asked them if they had any interaction difficulties in completing this part of the task. After a brief discussion about potential problems we introduced the *stereo view → tablet* and *tablet → stereo view* synchronization techniques as additional buttons on the tablet, previously hidden to the user. The experimenter demonstrated the use of this button by having the participant press the *stereo view → tablet* button to synchronize the TV and tablet views, unless the participant chose to use *tablet → stereo view*. With these additional features enabled and available for use throughout the continuing duration of the study, the user finalized their task at point B.

We then asked the participant to move to the opposite oblique point (C). At this point, the participant’s stereoscopic view of the data changed accordingly; however the tablet’s viewpoint again remained at its previous location. We asked the participant to manipulate the data such that someone standing at point A could see their original view from point A. Once again, we asked the participant if they had any interaction difficulties in completing this part of the task unless they voiced difficulties on their own. We also observed if they used one of the synchronization features; if so, we inquired why they chose that particular one.

### 4.3.3 Second Section

For the second section, we introduced the second interaction technique, following the same interface training procedure used for the first interaction technique, and switched to the second domain-specific dataset. Participants were asked to perform the same tasks as before. However, unlike in the first section, the *stereo view → tablet* and *tablet → stereo view* synchronization buttons were visible and available during the entire time. We reminded participants that they could use these buttons—and the gyroscope rotation—in addition to the second tablet interaction technique during the entire second section.

### 4.3.4 Questionnaire

After completing this second section, the participant completed a questionnaire and we conducted an in-person interview asking about their experiences with each interface and the interaction in

general. We asked participants about the effectiveness and intuitiveness of each technique (using Likert scales). We also asked about the amount of distraction they experienced with each interaction technique by looking back-and-forth between the tablet and stereoscopic display. We asked for their overall preference for one technique or the other, and the rationale for this choice. After the questionnaire, we interviewed participants to compare the pros and cons of each interface, and whether or not they felt limited in the types of interactions they could do with each setup. We asked specifically about the asynchronous tablet and stereoscopic displays, and whether or not this was an issue while interacting with the data. We additionally asked how they would extend either approach to directly address a particular visual analysis task.

#### 4.3.5 Data Recording and Procedure

We video-recorded the sessions and a secondary experimenter took observational notes. To ensure that we observed all possible permutations, we counter-balanced the order of the tBox and FI3D interfaces and the pairing of each interface with the two datasets. This also ensured that the initial reveal of the *stereo view → tablet* and *tablet → stereo view* synchronization features would occur across all possible dataset-interface combinations.

## 5 RESULTS AND OBSERVATIONS

In this section we summarize qualitative and descriptive quantitative results from our evaluation. We focus on how well the interaction modes and workflow address the re-synchronization of reference frames, and support intuitive 3D data analysis. We also discuss how the tablet interfaces' features support the interaction goals, and what may need to be taken into consideration when designing touch-based tablet interfaces for combined stereoscopic-monoscopic exploration of 3D datasets.

### 5.1 Different Reference Frames & their Synchronization

One of our main contributions is our discussion on how to address discrepancies between the monoscopic and stereoscopic frames (see Sec. 3.3). Here we discuss the participants' reaction to interacting with both a monoscopic tablet and a stereoscopic view as well as how they handled interaction after moving to a different location.

**Between Stereoscopic View and Monoscopic Tablet.** No participants reported problems or difficulties in looking back-and-forth between the tablet and the stereoscopic display. However, we observed participants using one of two types of behaviors when interacting with the tablet plus stereoscopic view setup:

- *tablet focus, stereo glance*: visual focus on the tablet view, only checking results on the stereo display; and
- *stereo focus, tablet glance*: despite initially looking at the tablet for interface cues, primary visual focus is the stereoscopic display with glances at the tablet for confirmation.

During the training with each interface, six participants focused primarily on their tablet, but soon shifted to a more stereo-focused view when presented with real datasets (M1, M2, M4, F2, F3, F4). While we did not track the focus of participants' gaze during our sessions, three of our participants (M4, F2, F3) rationalized their interface preference based on their ability to perform transformations without looking at the tablet (stereo focus), even using the tablet as a blind remote for periods of time. M4, for example, liked how the pinching gesture could be used to zoom without looking at the tablet, for either interface.

Two participants specifically cited their ability to interact with FI3D on the tablet while continuing to look at the stereoscopic display (M2, M4) when justifying their preference for the FI3D widget. F2—who indicated that he chose not to look at all at the tablet screen while interacting with the stereoscopic visualization—noted that rotations were simpler with the FI3D, as it does not require looking at the tablet at all times. The tBox, on the other hand, required some precision to operate and thus looking at the tablet. Preferring the FI3D because he could stay focused on the stereoscopic display, F3 thought the tBox was more difficult as he had “to check the tBox for where the edges or center is.”

**Resolving Reference Frames from New Positions.** During the first task, we were able to observe how participants reacted to the asynchronous reference frames of the tablet and their stereoscopic view. After moving to point B—but before introducing the synchronization—we noticed two participant behaviors. Four participants (M1, M4, F1, F3) realized the difference between the two views after a few moments of interaction. The participants cued into the difference when they noticed a mismatch between the  $z$ -axes of the tablet and stereoscopic view while performing rotations. The other four participants were so focused on the stereoscopic screen that they did not pay much attention to what was happening with the monoscopic tablet view. As a result, this group did not openly identify any unusual interactions. Finally, only F4 expressed any difficulties remembering the desired viewpoint found at the previous step—his efforts at point B were not in recalling the view, but simply recreating it.

After being introduced to the synchronization features and reaching position C, three participants used the *tablet → stereo view* button, and two participants used the *stereo view → tablet* button. The remaining three participants completed the task without using the buttons. During the second task, when synchronization features were enabled at all times, five participants used the *tablet → stereo view* button at position B; the remaining three participants completed the task without using the buttons. At position C in the second task, three participants used *stereo view → tablet* and three participants used *tablet → stereo view*. The remaining two participants continued without using the buttons. In the post-study interviews, two participants (F1, F3) noted that they liked being able to restore their previous view from the new position. Specifically during the task at position C—recreating the view for someone standing at position A—F1 thought it was better to have the view that he wanted at his current position and then rotate it into place.

All participants thought that it was important to have two separate views on the tablet and on the stereoscopic display; this is primarily due to the fact that the *tablet → stereo view* and *stereo view → tablet* synchronization techniques allowed participants to preserve and reuse their previous point of view from the tablet as they moved throughout the room. F3 noted that, while it is not useful to have duplicate views on both the tablet and the stereoscopic display, it is very useful to be able to change to another point of view.

### 5.2 Intuitiveness of the 3D Data Exploration

In our evaluation, we were also interested in the overall effectiveness of using touch-based interactions on monoscopic tablets in order to navigate stereoscopic views of 3D data. We gathered participant responses to Likert-scale questions asking about their ability to achieve specific views, their ability to do what they wanted to do and quickly achieve their goals. For both the tBox and FI3D

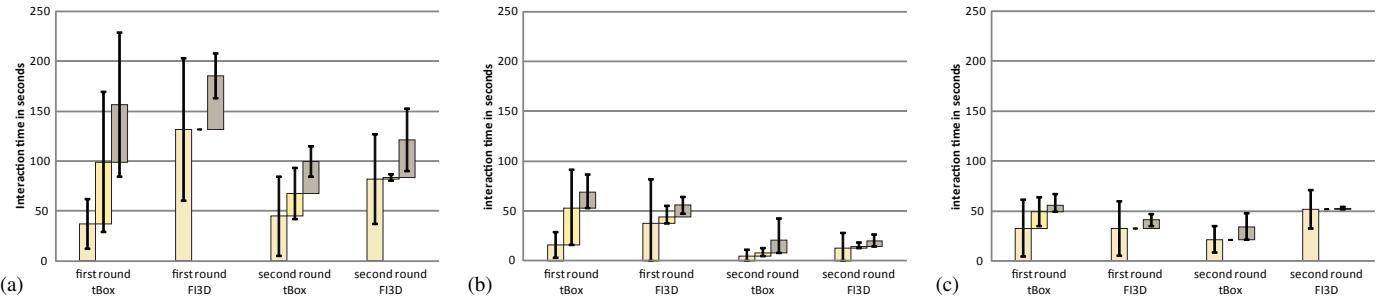


Fig. 6. Interaction times in the tablet interaction modes at points A (a), B (b), and C (c). In each graph, the bottom-left bar (light beige) shows rotations, the middle bar (yellow) shows gyroscope input, and the top-right bar (gray) shows other interactions. Error bars show 95% confidence intervals.

TABLE 2  
Results from Likert scale rating questions, values ranging from 1 (I completely disagree) to 5 (I completely agree).

	question	factor	median	SD
I could do what I wanted to do.	FI3D	4	0.52	
	tBox	4	0.89	
I could achieve my goals quickly.	FI3D	4	0.53	
	tBox	4	1.60	
It was clear to me how to achieve certain views.	FI3D	4	0.71	
	tBox	4	0.71	
It could be used without much explanation.	FI3D	4	0.83	
	tBox	4	1.16	
It required a lot of mental effort to use.	FI3D	2	0.74	
	tBox	2	0.93	

interfaces, participants agreed that they were able to accomplish what they wanted to do, they could quickly achieve their goals, and that it was clear how to achieve certain views, and use the system without much explanation (see Table 2).

As expected, we observed that the participants performed tasks faster the second time they tried to reproduce the view they had created in the previous step (see Fig. 6). Five of them quickly finished the task by using the *tablet → stereo view* button. Comparing the time spent at that phase during the first round (mean: 62.20 sec) and second round at that phase, a Levene's test did not show a violation of the homogeneity of variances ( $F(1, 14) = 1.365, p = 0.252 > 0.05$ ) and then an ANOVA proved the significant difference ( $F(1, 14) = 4.909, p = 0.044 < 0.05$ , effect size  $\eta^2 = 0.260$  with a confidence interval of [0, 0.536]).

Another concern was how our interaction framework would hold up to different datasets; we used two real-world datasets for each domain (four datasets total), and one participant from fluid mechanics noted that one of the flow datasets seemed more complicated than the other. We conducted an ANOVA to look for any potential effect that the dataset could have on the interaction times during the test. As shown in Table 3, we did not see a noticeable effect on the interaction time among the various datasets.

### 5.3 Contrasting Tablet Touch Interfaces

We implemented two tablet interfaces so as to demonstrate a range of tablet interfaces that could be applied to this interaction model. We did not find a statistically significant difference in the amount of time participants spent interacting with the two techniques Table 4. However, by understanding people's reactions to each interface, we can gain a sense of how the selection of the tablet interface may affect how people interact with and use the stereoscopic view.

TABLE 3  
Results of statistical analysis of effect of domain-specific datasets on amount of interaction time.

viewpoint	Levene's test		ANOVA	
	F(1, 14)	p	F(1, 14)	p
A	1.278	0.278	4.56	0.051
B	0.019	0.964	2.435	0.141
C	0.234	0.636	0.819	0.381

**tBox.** Only two participants preferred the tBox interface. Because the tBox's orientation is coupled to the dataset's orientation, several people had difficulty articulating transformations relative to their view. As a result, two participants wanted to reset the axis alignment of the tBox to correspond to their current perspective (M1, M3). Participants also perceived that the tBox had an implicit order of operations—"It is easier to rotate an object while it's inside of the box before translating it away" (F1). The precision of the tBox interface was interpreted both positively (M3 preferred the tBox because of its precision) and negatively (F3 felt that the tBox was "too precise").

**FI3D.** The remaining six participants preferred the FI3D interface. Two participants noted that the FI3D interface was particularly helpful when engaging in a free exploration of the data (M2, M3). Unlike the tBox interface, the order of data transformations (rotation, translation, and scaling) did not seem to matter as much (F2). FI3D was also noted for its ability to be used without looking at the tablet itself. One participant from structural biology (M4) noted that the FI3D interface had similarities with the interaction in PyMol, a visualization software package commonly used for 3D molecular data. However, one participant (F1) had difficulty using the FI3D interface, particularly when initiating a z-rotation or horizontal and vertical translations as both transformations are initiated using the same widget element. F1 also felt that FI3D combined too many separate visual elements that each interacted with different types of transformations. As a result, he preferred that the tBox represented just one controller, where all focus could be centered (F1).

**Gyroscope.** While gyroscopic navigation has frequently been used for remote navigation of 3D datasets in the past, our participants did not necessarily use the gyroscopic as a primary way of navigating through data when presented with a touch-based alternative. Two participants remarked that the gyroscope could be used for both small, precise rotational adjustments (M2, M3) as well as larger rotations (M3) to quickly change the point of view. M3 particularly felt that the gyroscope enabled easier access to free rotations and data exploration.

TABLE 4  
Evaluation of potential effects of the differences in interaction times between tBox and FI3D.

viewpoint	Levene's test		ANOVA	
	F(1, 14)	p	F(1, 14)	p
A	0.0038	0.9518	0.4134	0.531
B	0.0107	0.919	0.102	0.754
C	0.234	0.636	0.312	0.585

We also noted that our participants tended to rely more on the gyroscope for rotations when they navigated the visualization using the tBox, as opposed to the FI3D technique (see Fig. 6). One possible explanation is that the gyroscope provides a faster way than the tBox for simultaneously doing rotations using more than one axis. We ran a t-test, comparing the amount of interaction time with the gyroscope while using the tBox and the FI3D widget. We found a significant difference ( $t(7) = -2.502$ ,  $p = 0.041 < 0.05$ ; Cohen's  $d = 0.0844 \in [0.035, 1.692]$ ), the time spent using the gyroscope, with 0.724% during the tBox session and 29.730% during the FI3D session.

#### 5.4 Suggested Extensions

Participants had many suggested modifications and extensions of the tablet-stereo interface, particularly in re-imagining alternate uses for an asynchronous second screen. They relate, in particular, to reference frame mapping and the interaction design for combined mobile monoscopic with stationary stereoscopic visualization settings we report some of them next.

**Alternate roles for the monoscopic display.** Possible uses included storing views (M3, F4, F3), particularly for sharing with collaborators (M1, F4). Another participant (F1) wanted to entirely disconnect the tablet from the stereoscopic display and manipulate the tablet view without transforming the stereoscopic display. Another suggestion (F4) was to use the pair of displays to separately show “focus+context” views—e. g., a zoomed-in view of a molecule on one display, with the full molecule and an icon indicating camera position on the other. M3 suggested using the tablet as a “minimap” of the dataset, with the main view on the stereo screen. Given that we have addressed reference frame mapping between the tablet and the stereoscopic display, these new alternate roles for the monoscopic display are not only possible, but relatively straightforward functional extensions.

**Navigation.** Participants also wanted to be able to center the visualization around a defined area of interest (F4) by, e. g., introducing cutting planes (M1). Three participants attempted to use the gyroscope mode to do z-translation (M4, F1, F3).

**New Interactions.** Finally, one participant suggested using a music stand to hold the tablet while interacting with the stereoscopic display (F3). While this would limit their mobility in the visualization space somewhat, it would enable them to bi-manually interact with the tablet.

## 6 DISCUSSION AND CONCLUSION

Our main contribution in this paper is a detailed reflection on the different interaction modes for tablet-based navigation of a stereoscopic visualization and the transitions between them. While related setups exist for augmented reality scenarios, to the best of our knowledge our article is the first in-depth analysis that focuses specifically on the unique constraints of 3D visualization

applications. We based our investigation on a two-display setup that shows data and interaction elements also on the tablet. The reason for this design is that, ultimately, we want to support complex data analysis scenarios that require rich, potentially data-specific interaction techniques that go beyond simple 3D navigation.

For this purpose we discussed how reference frames may differ between the monoscopic view presented on the tablet and the interacting person’s stereoscopic view of the 3D dataset. In particular, we explained how these divergent reference frames are not problematic as long as the tablet’s virtual camera and the viewer’s stereoscopic view are approximately the same—even if tablet is held at a relaxed position with a “frozen” view. Our arguments are supported by evidence from research literature in sensorimotor adaptation and from our qualitative observational study—no participant mentioned problems with the tablet’s physical location being offset from the location of its virtual camera. Our participants were immersed in their perception of and interaction with the 3D data.

Moreover, with the help of synchronization tools, our participants were glad that the tablet’s view of the data was not constantly changing based on their position. In fact, most participants said they preferred separate views because, for example, it may allow them to save a view to which they could later return. We believe that this stability enabled people to more effectively control the stereoscopic view blindly, as the role of the visualization on the tablet becomes a stable yet understandable framework for navigation. Since our observational study was (intentionally, [36]) not designed to be a controlled experiment, we rely on other future controlled experiments to verify specific aspects such as this one with statistically valid results.

While our interaction design and its qualitative evaluation are applicable to 3D interaction and VR/AR in general, what makes this application domain unique is that, in the case of 3D visualization, both views are entirely virtual. We were able to avoid the problematic video camera-like interaction common in handheld AR applications altogether, and arrived at an interaction design that is free of awkward tablet handling or unclear reference frame mappings. Our discussion is thus **fundamental to future visualization systems** because they will increasingly rely on mobile devices as well as their integration with shared, stereoscopic environments. Moreover, as we focused specifically on the case in which the data is displayed on both the touch-enabled tablet and the large immersive display, our specific two-display interaction design opens up possibilities to add additional interactive data exploration capabilities on the tablet screen which would otherwise be difficult to impossible to realize if the hand-held controller would be an always eyes-free, touchpad-like input-only device.

Despite the fact that the specific implementation and comparison of tBox and FI3D is not our main contribution, our study still provided insight on the applicability of these specific interaction techniques to our specific setup. We observed that all participants could navigate the datasets using tBox and FI3D with the brief training we provided—with more training it would likely be even more effortless. Our study qualitatively compared these interface techniques, demonstrating their different strengths and weaknesses. Despite the majority of the participants favoring FI3D in our setting, this did not mean that it is better suited to 3D data navigation than tBox, per se. In fact, our participants pointed out that FI3D was better suited for free exploration, while tBox offered more precise control—both essential in visualization applications (G6 and [43]). Therefore, we believe that a combination of both techniques would be best in practice—possibly one as the default navigation, with

the other invoked using a spring-loaded mode [65], similar to how we implemented gyroscopic interaction.

Of course, our work also has a number of limitations that offer opportunities for future research. For example, we only worked with researchers from two domains—in which the data has unique qualities (e.g., flow typically in one direction from left to right, molecules with a [user-specified] up direction). It would be important to observe experts from other domains whose datasets have different characteristics, and who thus introduce new preferences on how to interact with such data. Also, we observed participants in a relatively small VR setting, and with a fairly constrained setup. It would be interesting to compare our findings against extended observations in larger and higher-quality VR environments (CAVEs or large walls) in which people have more freedom to move, use more elaborate visualization tasks/datasets, and provide advanced visualization techniques. Longitudinal studies would also provide more insight on the practical applicability of the proposed setup and highlight new ways to extend it, such as with other tablet-based techniques. It is also important to notice that our interaction design uses neither static nor dynamic peephole navigation [54]—we always show roughly the same scene (likely the whole dataset) in both views. It would be interesting to explore if the interaction needs to change if the VR setting was larger and users would carry out tasks that require zooming—our currently employed interaction techniques are more akin to the static peephole metaphor which appears to be inferior to the dynamic one for 2D interaction [54].

Naturally, our work also poses several new questions. One is how people handle both static and constantly updated views on the tablet. Our argumentation in this article was that it is beneficial to keep a frozen view on the tablet and only synchronize the tablet’s view to the immersive view once the user has moved far enough. The motivation for this design is that, as long as the person exploring a visualization keeps roughly at the same location, the mental mapping is not affected. At the same time, the user benefits from precise control due to the static frame of references and only needs to synchronize occasionally. However, our model does not keep users from constantly employing the *stereo view → tablet* (either by actively pressing the button or through a system mode)—at the expense of a lack of control as we hypothesize.

We would also be interested in how to design interactions for collaborative visualization settings, and how the presence of several people with different reference frames would affect the interaction design. Also, we are very interested in understanding how the tablet can be used to carry private data views, control additional visualization parameters, or show a differently parametrized/abstracted view. Innovations in display hardware such as tablets with localized tactile feedback [38], tangibly enhanced devices [46], or even deformable displays [67] will likely change how people can use them in an eyes-free way to control a large, stereoscopic visualization view, and would thus open exciting new possibilities for 3D data exploration and analysis.

## ACKNOWLEDGMENTS

We thank all our participants for their time and for providing the used datasets as well as Niklas Röber for a flow dataset. We also would like to thank Daniel F. Keefe for providing code used in our implementation, Douglas W. Cunningham for discussions on sensorimotor adaptation, and Mathieu Le Goc for help with the video. This work was supported, in part, by the French DigiTeo grant # 7588.

## REFERENCES

- [1] S. Abeele and O. Bock. Mechanisms for sensorimotor adaptation to rotated visual input. *Experimental Brain Research*, 139(2):248–253, 2001. doi>10.1007/s002210100768
- [2] S. Abeele and O. Bock. Sensorimotor adaptation to rotated visual input: Different mechanisms for small versus large rotations. *Experimental Brain Research*, 140(4):407–410, 2001. doi>10.1007/s002210100846
- [3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, Nov. 2001. doi>10.1109/38963459
- [4] F. Bedford. Perceptual learning. In D. L. Medin, ed., *The Psychology of Learning and Motivation*, volume 30, pp. 1–60. Academic Press, Inc., New York, 1993. doi>10.1016/S0079-7421(08)60293-5
- [5] F. L. Bedford. Of computer mice and men. *Cahiers de Psychologie Cognitives / Current Psychology of Cognition*, 13(4):405–426, 1994.
- [6] H. Benko, A. D. Wilson, and P. Baudisch. Precise selection techniques for multi-touch screens. In *Proc. CHI*, pp. 1263–1272. ACM, New York, 2006. doi>10.1145/1124772.1124963
- [7] A. Benzina, A. Dey, M. Tönnis, and G. Klinker. Empirical evaluation of mapping functions for navigation in virtual reality using phones with integrated sensors. In *Proc. APCHI*, pp. 149–158. ACM, New York, 2012. doi>10.1145/2350046.2350078
- [8] O. Bock. Basic principles of sensorimotor adaptation to different distortions with different effectors and movement types: A review and synthesis of behavioral findings. *Frontiers in Human Neuroscience*, 7(81), 2013. doi>10.3389/fnhum.2013.00081
- [9] M.-A. Bogorin, J. Luderschmidt, R. Dörner, and C. Geiger. ViNet – Interaction with information visualizations in VR applications via multi-touch and tangible user interfaces. In *Proc. 6<sup>th</sup> GI Workshop “Virtuelle und Erweiterte Realität”*, pp. 211–222. Shaker-Verlag, Aachen, 2009.
- [10] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: Mobile interaction through video. In *Proc. CHI*, pp. 2287–2296. ACM, New York, 2010. doi>10.1145/1753326.1753671
- [11] D. A. Bowman, E. Kruijff, J. J. LaViola, Jr., and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, Boston, 2005.
- [12] G. Bruder, F. Steinicke, and W. Stürzlinger. Touching the void revisited: Analyses of touch behavior on and above tabletop surfaces. In *Proc. INTERACT*, pp. 278–296. Springer, Berlin, 2013. doi>10.1007/978-3-642-40483-2\_19
- [13] G. Bruder, F. Steinicke, and W. Stürzlinger. To touch or not to touch? Comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces. In *Proc. SUI*, pp. 9–16. ACM, New York, 2013. doi>10.1145/2491367.2491369
- [14] S. Bryson. Virtual reality in scientific visualization. *Communications of the ACM*, 39(5):62–71, 1996. doi>10.1145/229459.229467
- [15] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1):341–377, 2011. doi>10.1007/s11042-010-0660-6
- [16] L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu, and Y.-P. Hung. Touching the void: Direct-touch interaction for intangible displays. In *Proc. CHI*, pp. 2625–2634. ACM, New York, 2010. doi>10.1145/1753326.1753725
- [17] D. Coffey, F. Korsakov, and D. F. Keefe. Low cost VR meets low cost multi-touch. In *Proc. ISVC*, volume 2, pp. 351–360. Springer, Berlin, Heidelberg, 2010. doi>10.1007/978-3-642-17274-8\_35
- [18] D. Coffey, N. Malbraaten, T. Le, I. Borazjani, F. Sotiropoulos, A. G. Erdman, and D. F. Keefe. Interactive slice WIM: Navigating and interrogating volume datasets using a multi-surface, multi-touch VR interface. *IEEE Transactions on Visualization and Computer Graphics*, 18(10):1614–1626, 2012. doi>10.1109/TVCG.2011.283
- [19] A. Cohé, F. Dècle, and M. Hatchet. tBox: A 3D transformation widget designed for touch-screens. In *Proc. CHI*, pp. 3005–3008. ACM, New York, 2011. doi>10.1145/1978942.1979387
- [20] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart. The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–72, June 1992. doi>10.1145/129882.129892
- [21] H. A. Cunningham. Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3):493–506, 1989. doi>10.1037/0096-1523.15.3.493
- [22] F. Daiber, L. Li, and A. Krüger. Designing gestures for mobile 3D gaming. In *Proc. MUM*, pp. 3:1–3:4. ACM, New York, 2012. doi>10.1145/2406367.2406371
- [23] M. Deering. High resolution virtual reality. *SIGGRAPH Computer Graphics*, 26(2):195–202, 1992. doi>10.1145/142920.134039

- [24] M. Deller and A. Ebert. ModControl – Mobile phones as a versatile interaction device for large screen applications. In *Proc. INTERACT*, pp. 289–296. Springer, Berlin, 2011. doi> 10.1007/978-3-642-23771-3\_22
- [25] Y. Du, H. Ren, G. Pan, and S. Li. Tilt & touch: Mobile phone for 3D interaction. In *Proc. Ubicomp*, pp. 485–486. ACM, New York, 2011. doi> 10.1145/2030122.2050183
- [26] J. Grubert, R. Grasset, and G. Reitmayr. Exploring the design of hybrid interfaces for augmented posters in public spaces. In *Proc. NordiCHI*, pp. 238–246. ACM, New York, 2012. doi> 10.1145/2399016.2399053
- [27] M. Hatchet, B. Bossavit, A. Cohé, and J.-B. de la Rivière. Toucheo: Multitouch and stereo combined in a seamless workspace. In *Proc. UIST*, pp. 587–592. ACM, New York, 2011. doi> 10.1145/2047196.2047273
- [28] M. Hancock, T. ten Cate, and S. Carpendale. Sticky tools: Full 6DOF force-based interaction for multi-touch tables. In *Proc. ITS*, pp. 145–152. ACM, New York, 2009. doi> 10.1145/1731903.1731930
- [29] C. Hand. A survey of 3D interaction techniques. *Computer Graphics Forum*, 16(5):269–281, 1997. doi> 10.1111/1467-8659.00194
- [30] R. Hardy and E. Rukzio. Touch & interact: Touch-based interaction of mobile phones with displays. In *Proc. MobileHCI*, pp. 245–254. ACM, New York, 2008. doi> 10.1145/1409240.1409267
- [31] I. Haug. Designing for expert users. Blog post, <http://www.usablemachine.com/blog/designing-for-expert-users/>, Apr. 10, 2012. Visited Nov. 2014.
- [32] P. Isenberg and T. Isenberg. Visualization on interactive surfaces: A research overview. *i-com*, 12(3):10–17, 2013. doi> 10.1524/com.2013.0020
- [33] P. Isenberg, T. Isenberg, T. Hessermann, B. Lee, U. von Zadow, and A. Tang. Data visualization on interactive surfaces: A research agenda. *IEEE Computer Graphics and Applications*, 33(2):16–24, 2013. doi> 10.1109/MCG.2013.24
- [34] T. Isenberg. Position paper: Touch interaction in scientific visualization. In *Proc. DEXIS (ITS Wksp.)*, pp. 24–27, 2011.
- [35] T. Isenberg and M. Hancock. Gestures vs. postures: ‘Gestural’ touch interaction in 3D environments. In *Proc. 3DCHI (CHI Workshop)*, pp. 53–61, 2012.
- [36] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013. doi> 10.1109/TVCG.2013.126
- [37] J. Jankowski and M. Hatchet. A survey of interaction techniques for interactive 3D environments. In *Eurographics State of the Art Reports*, pp. 65–93. Eurographics Association, Goslar, Germany, 2013. doi> 10.2312/conf/EG2013/stars/065-093
- [38] Y. Jansen, T. Karrer, and J. Borchers. MudPad: Tactile feedback and haptic texture overlay for touch surfaces. In *Proc. ITS*, pp. 11–14. ACM, New York, 2010. doi> 10.1145/1936652.1936655
- [39] N. Katzakis, K. Kiyokawa, M. Hori, and H. Takemura. Plane-casting: 3D cursor control with a smartphone. In *Proc. 3DCHI (CHI Workshop)*, pp. 13–22. ACM, New York, 2012.
- [40] D. F. Keefe. Integrating visualization and interaction research to improve scientific workflows. *IEEE Computer Graphics and Applications*, 30(2):8–13, Mar./Apr. 2010. doi> 10.1109/MCG.2010.30
- [41] D. F. Keefe and T. Isenberg. Reimagining the scientific visualization interaction paradigm. *IEEE Computer*, 46(5):51–57, 2013. doi> 10.1109/MC.2013.178
- [42] J.-S. Kim, D. Gračanin, K. Matković, and F. Quek. iPhone/iPod Touch as input devices for navigation in immersive virtual environments. In *Proc. VR*, pp. 261–262. IEEE, Piscataway, NJ, 2009. doi> 10.1109/VR.2009.4811045
- [43] T. Klein, F. Guéniat, L. Pastur, F. Vernier, and T. Isenberg. A design study of direct-touch interaction for exploratory 3D scientific visualization. *Computer Graphics Forum*, 31(3):1225–1234, 2012. doi> 10.1111/j.1467-8659.2012.03115.x
- [44] S. Knoedel and M. Hatchet. Multi-touch RST in 2D and 3D spaces: Studying the impact of directness on user performance. In *Proc. 3DUI*, pp. 75–78. IEEE Computer Society, Los Alamitos, 2011. doi> 10.1109/3DUI.2011.5759220
- [45] W. Krüger and B. Fröhlich. The responsive workbench. *IEEE Computer Graphics and Applications*, 14(3):12–15, 1994. doi> 10.1109/38.279036
- [46] D. M. Krum, T. Phan, L. Cairco Dukes, P. Wang, and M. Bolas. Tablet-based interaction panels for immersive environments. In *Proc. VR*, pp. 91–92. IEEE, Piscataway, NJ, 2014. doi> 10.1109/VR.2014.6802066
- [47] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012. doi> 10.1109/TVCG.2011.279
- [48] D. Lee, J.-I. Hwang, G. J. Kim, and S. C. Ahn. 3D interaction using mobile device on 3D environments with large screen. In *Proc. MobileHCI*, pp. 575–580. ACM, New York, 2011. doi> 10.1145/2037373.2037463
- [49] G. A. Lee, U. Yang, Y. Kim, D. Jo, K.-H. Kim, J. H. Kim, and J. S. Choi. Freeze-Set-Go interaction method for handheld mobile augmented reality environments. In *Proc. VRST*, pp. 143–146. ACM, New York, 2009. doi> 10.1145/1643928.1643961
- [50] H.-N. Liang, J. Trenchard, M. Semegen, and P. Irani. An exploration of interaction styles in mobile devices for navigating 3D environments. In *Proc. APCHI*, pp. 309–313. ACM, New York, 2012. doi> 10.1145/2350046.2350062
- [51] H.-N. Liang, C. Williams, M. Semegen, W. Stuerzlinger, and P. Irani. An investigation of suitable interactions for 3D manipulation of distant objects through a mobile device. *International Journal of Innovative Computing, Information and Control*, 9(12):4737–4752, 2013.
- [52] J. F. Lucas and D. A. Bowman. Design and evaluation of 3D multiple object selection techniques. Report, Virginia Polytechnic Institute and State University, USA, 2005.
- [53] C. Lundström, T. Rydell, C. Forsell, A. Persson, and A. Ynnerman. Multi-touch table system for medical visualization: Application to orthopedic surgery planning. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1775–1784, Dec. 2011. doi> 10.1109/TVCG.2011.224
- [54] S. Mehra, P. Werkhoven, and M. Worring. Navigating on handheld displays: Dynamic versus static peephole navigation. *ACM Transactions on Computer-Human Interaction*, 13(4):448–457, Dec. 2006. doi> 10.1145/1188816.1188818
- [55] M. A. Nacenta, P. Baudisch, H. Benko, and A. Wilson. Separability of spatial manipulations in multi-touch interfaces. In *Proc. Graphics Interface*, pp. 175–182. CIPS, Toronto, 2009.
- [56] M. Novotný, J. Lacko, and M. Samuelčík. Applications of multi-touch augmented reality system in education and presentation of virtual heritage. *Procedia Computer Science*, 25:231–235, 2013. doi> 10.1016/j.procs.2013.11.028
- [57] V. Oblaender and M. Eibl. Study of interaction concepts in 3D virtual environment. In *Proc. HCI*, volume 4, pp. 706–711. Springer, Berlin, 2013. doi> 10.1007/978-3-642-39330-3\_76
- [58] T. Ohnishi, N. Katzakis, K. Kiyokawa, and H. Takemura. Virtual interaction surface: Decoupling of interaction and view dimensions for flexible indirect 3D interaction. In *Proc. 3DUI*, pp. 113–116. IEEE Computer Society, Los Alamitos, 2012. doi> 10.1109/3DUI.2012.6184194
- [59] A. Olwal and S. Feiner. Spatially aware handhelds for high-precision tangible interaction with large displays. In *Proc. TEI*, pp. 181–188. ACM, New York, 2009. doi> 10.1145/1517664.1517705
- [60] J. Rekimoto. Navicam: A magnifying glass approach to augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):399–412, Aug. 1997.
- [61] G. Robles-De-La-Torre. The importance of the sense of touch in virtual and real environments. *IEEE MultiMedia*, 13(3):24–30, 2006. doi> 10.1109/MMUL.2006.69
- [62] L. Schemali and E. Eisemann. Design and evaluation of mouse cursors in a stereoscopic desktop environment. In *Proc. 3DUI*, pp. 67–70. IEEE Computer Society, Los Alamitos, 2014. doi> 10.1109/3DUI.2014.6798844
- [63] D. Schmalstieg, L. M. Encarnaçao, and Z. Szalavári. Using transparent props for interaction with the virtual table. In *Proc. I3D*, pp. 147–153. ACM, New York, 1999. doi> 10.1145/300523.300542
- [64] D. Schmalstieg and D. Wagner. Experiences with handheld augmented reality. In *Proc. Mixed and Augmented Reality*, pp. 3–18. IEEE Computer Society, Los Alamitos, 2007. doi> 10.1109/ISMAR.2007.4538819
- [65] A. J. Sellen, G. P. Kurtenbach, and W. A. S. Buxton. The prevention of mode errors through sensory feedback. *Human Computer Interaction*, 7(2):141–164, 1992. doi> 10.1207/s15327051hci0702\_1
- [66] K. Shoemake. ARCBALL: A user interface for specifying three-dimensional orientation using a mouse. In *Proc. Graphics Interface*, pp. 151–156. Morgan Kaufmann, San Francisco, 1992.
- [67] J. Tsimeris, C. Dedman, M. Broughton, and T. Gedeon. ForceForm: A dynamically deformable interactive surface. In *Proc. ITS*, pp. 175–178. ACM, New York, 2013. doi> 10.1145/2512349.2512807
- [68] M. Uflacker and D. Busse. Complexity in enterprise applications vs. simplicity in user experience. In *Proc. HCI International*, pp. 778–787. Springer, Berlin, Heidelberg, 2007. doi> 10.1007/978-3-540-73111-5\_87
- [69] D. Valkov, A. Giesler, and K. H. Hinrichs. Evaluation of depth perception for touch interaction with stereoscopic rendered objects. In *Proc. ITS*, pp. 21–30. ACM, New York, 2012. doi> 10.1145/2396636.2396640
- [70] D. Valkov, A. Giesler, and K. H. Hinrichs. Imperceptible depth shifts for touch interaction with stereoscopic objects. In *Proc. CHI*, pp. 227–236. ACM, New York, 2014. doi> 10.1145/2556288.2557134
- [71] D. Valkov, F. Steinicke, G. Bruder, and K. Hinrichs. 2D touching of 3D stereoscopic objects. In *Proc. CHI*, pp. 1353–1362. ACM, New York, 2011. doi> 10.1145/1978942.1979142

- [72] D. Valkov, F. Steinicke, G. Bruder, K. H. Hinrichs, J. Schöning, F. Daiber, and A. Krüger. Touching floating objects in projection-based virtual reality environments. In *Proc. EGVE/EuroVR/VEC*, pp. 17–24. Eurographics Association, Goslar, Germany, 2010. doi>10.2312/EGVE/IVR10/017024
- [73] D. Wagner. *Handheld Augmented Reality*. PhD thesis, Graz University of Technology, Austria, Oct. 2007.
- [74] C. Ware and R. Arsenault. Frames of reference in virtual object rotation. In *Proc. APGV*, pp. 135–141. ACM, New York, 2004. doi>10.1145/1012551.1012576
- [75] D. Watson, M. Hancock, R. L. Mandryk, and M. Birk. Deconstructing the touch experience. In *Proc. ITS*, pp. 199–208. ACM, New York, 2013. doi>10.1145/2512349.2512819
- [76] R. B. Welch. *Perceptual Modification: Adapting to Altered Sensory Environments*. Academic Press, New York, 1978.
- [77] A. Wiebel, F. M. Vos, D. Foerster, and H.-C. Hege. WYSIWYP: What you see is what you pick. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2236–2244, Dec. 2012. doi>10.1109/TVCG.2012.292
- [78] L. Yu, K. Efstathiou, P. Isenberg, and T. Isenberg. Efficient structure-aware selection techniques for 3D point cloud visualizations with 2DOF input. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2245–2254, Dec. 2012. doi>10.1109/TVCG.2012.217
- [79] L. Yu, P. Svetachov, P. Isenberg, M. H. Everts, and T. Isenberg. FI3D: Direct-touch interaction for the exploration of 3D scientific visualization spaces. *IEEE TVCG*, 16(6):1613–1622, 2010. doi>10.1109/TVCG.2010.157
- [80] S. Zhai and P. Milgram. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proc. CHI*, pp. 320–327. ACM, New York, 1998. doi>10.1145/274644.274689



**Tobias Isenberg** is a senior research scientist with the AVIZ project team at INRIA-Saclay, France. He received his doctoral degree from the University of Magdeburg, Germany, in 2004. Previously he held positions as post-doctoral fellow at the University of Calgary, Canada, and as assistant professor at the University of Groningen, the Netherlands. His research interests comprise topics in scientific visualization, in illustrative and non-photorealistic rendering, and interactive visualization techniques.



**David López** is a MSc student of the Master in Engineering program at the University of Antioquia, Colombia, where he also received his engineering degree in computer sciences. He has worked in the Tele-Health field with mobile and custom hardware devices. He was a research intern in the AVIZ project team at INRIA-Saclay, France. He is interested in several topics of computer graphics, specifically global illumination, real-time rendering, and scientific visualization.



**Lora Oehlberg** is an assistant professor in the Computer Science department at the University of Calgary, Canada. She received her doctoral degree from the University of California at Berkeley, USA, in 2012. Previously she was a post-doctoral researcher with the InSitu project team at Inria-Saclay, France. Her research interests include human-computer interaction, particularly in technologies that support creativity, collaboration, and fabrication.



**Candemir Doger** is a MSc candidate in computer science program in Sabanci University, Turkey. He received his bachelor degree in computer science from Sabanci University, Turkey, in 2011. Previously he worked as a freelance developer on various projects in Turkey and as a research intern in AVIZ project team at INRIA-Saclay, France. He is also currently working as a software developer in SAP Innovation Center in Istanbul, Turkey. He is interested in information visualization and mobile platforms.