Bi-3D: Bi-Manual Pen-and-Touch Interaction for 3D Manipulation on Tablets

Ken Pfeuffer
Bundeswehr University, Germany
Aarhus University, Denmark
ken@cs.au.dk

Abdullatif Dinc
LMU Munich
Germany
dinc.abdullatif@campus.lmu.de

Jan Obernolte
LMU Munich
Germany
jan.obernolte@campus.lmu.de

Radiah Rivu
Bundeswehr University
Germany
radiah.rivu@unibw.de

Yasmeen Abdrabou
Bundeswehr University
Germany
yasmeen.essam@unibw.de

Franziska Shelter
LMU Munich
Germany
franziska.shelter@campus.lmu.de

Yomna Abdelrahman
Bundeswehr University
Germany
yomna.abdelrahman@unibw.de

Florian Alt
Bundeswehr University
Germany
florian.alt@unibw.de

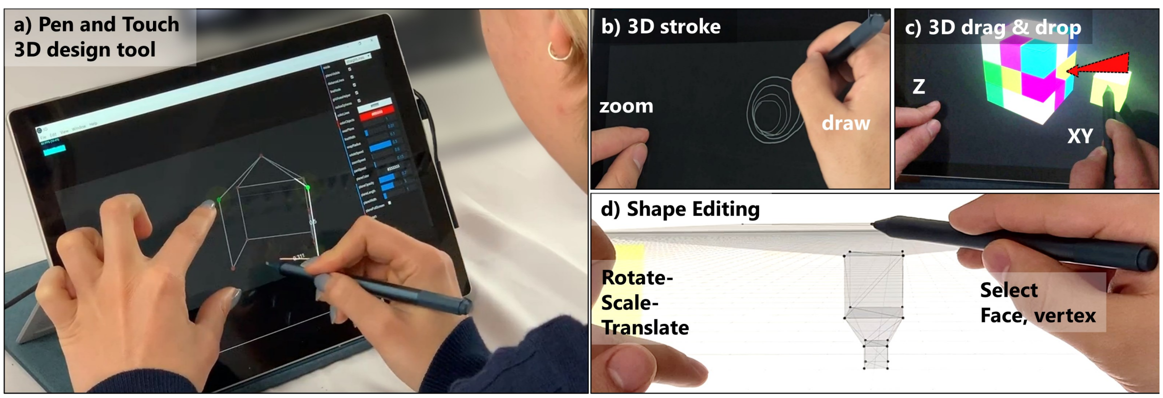

Figure 1: *Bi-3D* is an interaction concept to create and manipulate 3D objects and sketches. E.g., to edit shape points (a), draw a 3D stroke (b), 3D drag & drop an object (c), or scale shape (d) – enabled by a dynamic coupling of pen and touch in both hands.

ABSTRACT

Tablets are attractive for design work anywhere, but 3D manipulations are notoriously difficult. We explore how engaging the stylus and multi-touch in concert can render such tasks easier. We introduce Bi-3D, an interaction concept where touch gestures are combined with 2D pen commands for 3D manipulation. For example, for a fast and intuitive 3D drag & drop technique: the pen drags the object on-screen, and parallel pinch-to-zoom moves it in the third dimension. In this paper, we describe the Bi-3D design space, crossing two-handed input and the degrees-of-freedom (DOF) of 3D manipulation and navigation tasks. We demonstrate sketching and manipulation tools in a prototype 3D design application, where users can fluidly combine 3D operations through alternating and parallel use of the modalities. We evaluate the core technique, bi-manual 3DOF input, against widget and mid-air baselines in an object movement task. We find that Bi-3D is a fast and practical way for multi-dimensional manipulation of graphical objects, promising to facilitate 3D design on stylus and tablet devices.

UIST '21, October 10–14, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8635-7/21/10…$15.00
https://doi.org/10.1145/3472749.3474741

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**.

KEYWORDS

3D design, tablet computing, pen, multi-touch, bi-manual interface

ACM Reference Format:
Ken Pfeuffer, Abdullatif Dinc, Jan Obernolte, Radiah Rivu, Yasmeen Abdrabou, Franziska Shelter, Yomna Abdelrahman, and Florian Alt. 2021. Bi-3D: Bi-Manual Pen-and-Touch Interaction for 3D Manipulation on Tablets. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21), October 10–14, 2021, Virtual Event, USA*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3472749.3474741

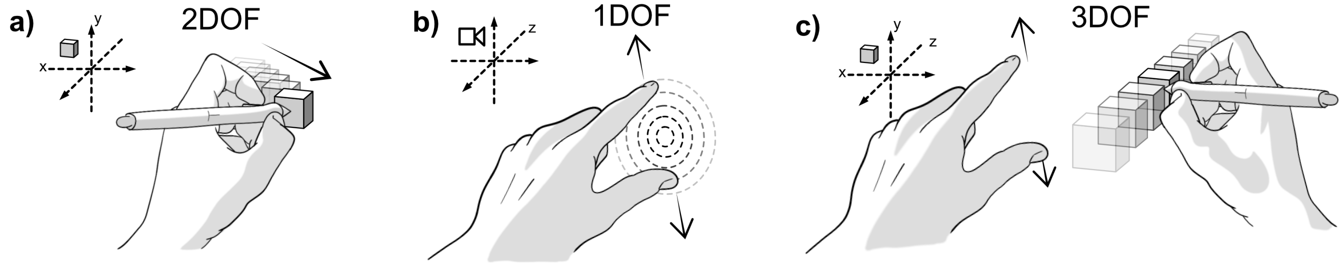

Figure 2: The Bi-3D concept exploits bi-manual gestures for extension to 3D object interaction. For example, individually, a system can support a 2D drag & drop via pen (a) and a 1D pinch-to-zoom gesture to change camera depth (b). In combination, they offer possibilities to extend to 3DOF operations, useful for 3D transformations such as 3D drag & drop (c).

1 INTRODUCTION

As the possibilities to experience 3D content are ever rising, so are the needs to support its creation across various stages of the design process [44]. Direct manipulation user interfaces (UI) – such as those supported by the Apple IPad, Microsoft Surface, and Samsung Galaxy Note – are intriguing for the design of 3D material as they afford a unique trade-off between ease of use, comfort, and precision. A major challenge herein lies in the establishment of screen-space models that map 2D inputs to virtual 3D object transformations.

Techniques for 2D–3D mappings can be classified into (1) augmenting the GUI by offering widgets, *separating the task* in lower-dimensional transformations, and (2) augmenting the input by supporting *integral mappings* where multiple inputs merge to expressive multiple degrees-of-freedom (DOF) interaction [25]. Current applications are based on the former approach, allowing one sub task to be precisely operated without affecting another [59, 62]. Yet, a frequent use of widgets can significantly impact and lengthen the workflow. Drag & drop for instance – performed with a single command in 2D contexts – is in 3D decomposed into multiple sub-steps even though we intuitively think of it as a the singular task.

We explore how a particular UI constellation, two-handed pen and touch, can advance the user's 3D manipulation capabilities. Styli offer unique affordances suitable for design work, such as input precision and minimal occlusion of content. Moreover, their compatibility with touch gestures is promising for composing multi-modal inputs that unify to higher-level commands [22]. Figure 1 shows several example applications of useful tools. For example, to draw a 3D spiral line, the user simply draws circles on the canvas while zooming out in parallel (b). Moving objects can be achieved with an adapted technique where the pen moves the object in 2D as expected, while a pinch gesture adjust the object's depth in the scene (c). In principle, many touch gestures can be unified with pen actions in many application contexts including shape editing (d). Those interactions become possible through a seamless interplay between the modalities and targeted operations, something that can significantly empower users but is currently barely supported.

Our conceptual point of departure is the question of how our beloved gestures, such as pan & zoom for touch and drag & drop for the pen, can be extended to 3D. In this case, the task structure – navigating for-/backward and dragging on-screen – perceptually aligns closely with the structure of a 3DOF interaction task. A bi-manual UI can be complemented with gestures that incorporate such structures (Figure 2). Uni-manually, users simply perform the expected interaction – either object dragging in 2D space (a) or zooming the canvas (b). In combination, this leads to the ability of 3DOF input (c).

We call this interaction concept *Bi-3D*: touch gestures that enhance the 2D pen inputs to form 3D commands to facilitate 3D manipulation. It can be useful for a range of applications, from extending 2D applications (e.g., Paint) with the third dimension to extending 3D tools with extra 3DOF input mechanisms.

In this paper, we first describe the Bi-3D design space of screen-space formulations and general approaches for 3D manipulations. We then present a prototype application, that demonstrates interaction techniques and how they could be used in a real application. For instance, how sketching and drawing can be enhanced to a third dimension, while retaining the simplicity of single-stroke pen operations. Or, how shape manipulations are enhanced by mechanisms to rapidly switch between translation, rotation, and scaling transformation through gesture moding. Lastly, we present an empirical evaluation to get insights into performance in an object docking task. Our findings provide promising insights: users are significantly faster compared to widgets, while no significant differences were found to an integral, mid-air pointing technique.

Taken together, our contributions are:

- The Bi-3D concept that enhances pen commands with touch gestures for a new category of 3DOF pen + touch techniques to render complex 3D tasks easier.
- A design space for screen-space mappings between the modalities and 3D-RST tasks, leading two base approaches:
  - *Navigation + Manipulation* as 3D-extended 2D operations.
  - *Manipulation x 2* as enhanced single-object manipulation.
- Techniques for applications that spawned from the design space, demonstrating the variety of its utility in:
  - Sketching tools where users can rapidly create 3D strokes.
  - Object manipulation tools where users can mix 3D-RST tasks.
- A study of 3DOF pen performance, yielding new insights on how a basic instantiation can surpass the current baseline and perform on-par with a mid-air technique.

2 RELATED WORK

Our work extends past Pen + Touch work with the first in-depth 3DOF investigation, relates to sketching and CAD tools, and extends prior 3D-Touch work by exploring combination with pen.

2.1 Bi-manual Pen + Touch

Bi-manual interfaces, where users control multiple degrees of freedom or sub-tasks with both hands, have been shown to be efficient for a range of tasks such as geometry manipulation, alignment of graphics, or interaction with several targets [9, 28, 31, 32]. What contributes to the efficiency is the ability to perform concurrent two-handed manipulation and a logical structure across the hands. For example, *asymmetric* dependent tasks benefit when aligned to Guiard's Kinematic Chain (KC) model of bi-manual interaction [9, 18]. This is grounded in principles of *spatial/temporal reference*, where the non-dominant hand (NDH) sets the spatial frame of reference of and precedes the dominant hand (DH)'s actions, and *asymmetric scale of motion* where the hands operate in asymmetric spatio-temporal scale. Pen and touch can be used in this way in their i) alternating use, e.g., the user first scrolls the canvas, and then sketches with the pen, and ii) simultaneous use, e.g., when users hold a note object and cut it with the pen [22].

Another bi-manual category are *more symmetric* constellations. For instance, in Tape Drawing [2], as well as several other systems [41, 67, 69], the dominant hand precedes and sets the spatial frame of reference. In principle, this is relatively natural, considering how we use our hands with nail and hammer or bow and arrow. Pen and touch can be extended in this way: users first employ the pen, then users employ touch gestures in addition. Such *touch-activated pen (TAP) gestures* [69] can enrich the pen's basic functionality with further modes. For example, for whitespace generation [67, 69], 2D selections [69], or on-the-fly stroke adjustments [40]. In our work, we aim to include both bi-manual categories, as they are complementary, to support the more complex 3D manipulations.

Researchers extensively explored the application space of pen and touch in both hands. Brandl et al. firstly revealed the design space of combining both modalities to form new interaction techniques [7]. A seminal paper in this space is Pen + Touch = New Tools [22], that firstly provided a coherent set of techniques for note-taking and scrapbooking of materials [22]. Such a coherent set allows to avoid menus and mode switching time. In the more complex 3D domain of our work, we explore two coherent sets and expand their feature range with menus. Across many graphical applications, several papers have suggested a default UI of pan & zoom gestures for UI navigation with a pen for commands is useful for sketching [21], visualisation explorations [54], document editing [51, 68], or vector graphics work [7, 66]. The Thumb + Pen concept extended it to mobile tablets when holding the device with one hand [49]. Cami et al. explored the unimanual pen+touch input space, leveraging the touch information when the hand that holds the pen is on the table [10]. Xia et al. proposed the ZoomCatcher, as a context sensitive UI widget to enhance sketching on large screens, demonstrating the synergy of simultaneous pen + pinch-to-zoom gesture [63]. A few papers have also proposed pen (and touch) operated object layering [13, 58] and line-connecting [48], pointing to a promising avenue of research toward 3D interaction.

2.2 3D Sketching

Several efforts pursued to extend sketching to 3D forms. This includes work on approximately converting strokes as 3D lines [1, 17, 24, 65], interpreting multiple strokes as single 3D lines [12], and using a mechanically collapsible pen [33, 61] or tablet [16, 64]. A practical approach is combined pen and touch, as studied by Lopes et al. for 3D modelling [35]. Aiming for comfort and familiarity, two-handed interactions to manipulate camera operations while using a pen to sketch can indeed work well. Their study revealed higher bi-manual user performance than an interface operated purely by a pen. Such an interaction model has been adopted to sketch storyboards (StoreoBoard [21]), where users perform nondominant hand touches to change between different layers of the board. A more dynamic switch is supported by MentalCanvas, where users sketch on 2D planes defined by the user with the pen, while navigation operations allows the user to rapidly traverse the 3D space to view and operate on the planes [15, 56]. We extend this application-oriented avenue with a focus on pen and touch for 3D interactions.

2.3 Commercial CAD Software

A set of commercial tools exist for a pen and touch UI, such as Autodesk Maya, Blender, Apple Note 3D, Trimple Sketchup, or Shapr3D. These tools support a interaction model of pan & zoom and pen commands, aligning with existing research [22, 35]. The majority of the tools are based on the traditional widgets to enable 3D interaction based on separated DOF. The simultaneous use of the modalities is marginally adopted (e.g., holding a virtual ruler with touch to draw straight lines with a pen). We complement those works, focusing on how the modalities can exploit a more seamless coupling as an enabler for 3DOF interactions.

2.4 3D-Touch

A body of work investigated direct manipulation for 3D objects (see Mendes et al.'s survey for details [42]). Techniques can be categorised based on Jacob's work on integrality of input devices [25]. An integral device supports movement across multiple dimensions of control (DOF integration), whereas a separable device restricts movement along one dimension at a time (DOF separation). A common type of DOF separation in commercial applications are widgets that provide handles for each DOF, which we will revisit as a study baseline. DOF integration is part of the proposed techniques.

Some approaches map the pinch-to-zoom gesture to control an object's depth as it is appealing that an object appears larger the closer it is to the scene camera [8, 20, 34, 50]. For example, Sticky Tools [20], and Reisman et al.'s screen-space formulation [50] use single and two-finger dragging to move an object in screen space, while the distance of two fingers manipulates the depth. As indirect mappings can improve efficiency and precision [29], the two-finger pinch operation can be performed on widgets offset from the object as in Balloon Selection [6] or Triangle Cursor [55]. The Z-technique [37] and DS3 [38] provide a variant: the first touch selects and drags an object in screen space, and second touch dragging anywhere on the screen manipulates depth – a variant that fared well in their evaluation. We extend prior art with exploring how touch gestures support the stylus based multi-DOF object manipulation.

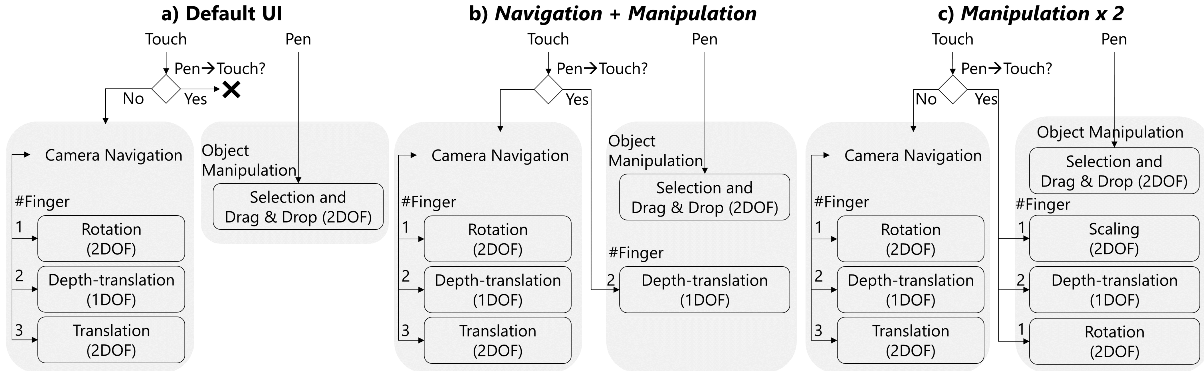

Figure 3: High-level input state models. (a) A simplified model of a typical 3D UI with separable pen and touch, and two new models where touch input enhances the pen's manipulation with (b) and without (c) parallel navigation operations.

3 BI-3D

The literature review showed a rich history of past achievements in the two spaces of 2D pen and touch and 3D interaction. Bi-3D lies at the intersection, as a concept that enables interaction techniques that exploit the two modalities for 3D manipulations. This necessitates the design of screen-space formulations [50], to map 2D input devices to 3D content, that we elaborate in this section. We first describe a typical UI where multi-touch and a stylus are used in a 3D tool without simultaneous use. We then describe potential formulations when combining the two-handed input sensed with the structures of 3D manipulations. We also analyse how the input possibilities can be integrated complementary to default 3D navigation and manipulation tasks, with and without visual feedback, before we present applications.

3.1 Default UI Input States

Figure 3a illustrates the input states that the system can support for camera navigation and object manipulation. Depending on how they employ the pen and fingers on the screen, users can switch between the following operations.

3.1.1 Navigation (Touch). camera navigation is intuitive with multi-touch gestures and screen-space formulations have been addressed in prior work [19]. Here we describe one potential instance that aims for simplicity, and for different approaches we refer to the literature – our main focus is afterwards on the combinations with the pen. Our set separates transformations by one, two, and three finger touch gestures to support rotation, panning in depth, and panning in screen-space, respectively.

- 1-touch rotation: The rotation reference is set as a point on the camera forward vector (adjustable in settings). Single-finger drag gestures rotate the camera around this point.
- 2-touch depth translation: For depth movement, we map the Euclidean distance between the two pinch fingers to the third dimension parameter. The distance is mapped 1:1 to screen space to remain consistent with the 2D panning. E.g., if a pinch gesture reduces the distance between both fingers from 100 px to 50 px, the camera moves by 50 px in depth.
- 3-touch panning: With three fingers on the screen, their parallel movement in screen-space pans the camera accordingly.

Note that the gestures are in principle interchangeable, e.g., if a user prefers to pan with 1-touch, to align with 2D pan & zoom, it can be swapped in system configurations.

3.1.2 Manipulation (Pen). Using the pen for object manipulation is straightforward: select an object with the pen tip and begin a 2DOF operation, e.g., drag & drop. This is different to Lopes et al.'s system who do not use pen dragging [35]. Target selection by the 2D pen in a 3D scene can be implemented (1) by a common 2D–3D projection. A ray is cast from the pen contact position, perpendicular to the screen plane. The intersecting object is selected, and users can start manipulations. However, this approach is difficult for making freeform sketches in empty 3D space. (2) Alternatively, it can be based on 3D sketching. This is inspired by how digital 2D sketching happens like on pen and paper, in the screen-space plane (e.g., Microsoft Paint or Adobe Photoshop). This can be extended to 3D space, when considering a 2D surface as a 2D slice of the 3D world [52]. The system can help the user to generate and move the surface [15, 26], but also use a fixed plane in screen-space for simplicity [21, 30, 64]. We include the latter approach, where the drawing plane is at the screen-space, which users can navigate by the camera navigation gestures. This allows users to navigate to an arbitrary slice of the 3D scene and start a line freely in 3D space.

3.2 Bi-3D: From 2D to 3D Manipulation

Building on the default input states, Bi-3D is an interaction concept that extends the pen's 2D operations to 3D in three steps:

1. The user first selects the object in the 3D scene by the pen. 2D operations can be conducted.
2. The user can contact their finger on the screen for advanced manipulation of the object under the pen tip. The finger controls an extra operation on top of the pen's current operation.
3. Moving the pen in parallel results in a 3DOF object manipulation.

Releasing the touch inputs returns the user to 2D pen commands. If the stylus is contacted after a touch, it is discounted as a simultaneous operation as the touch gestures already initiated a navigation.

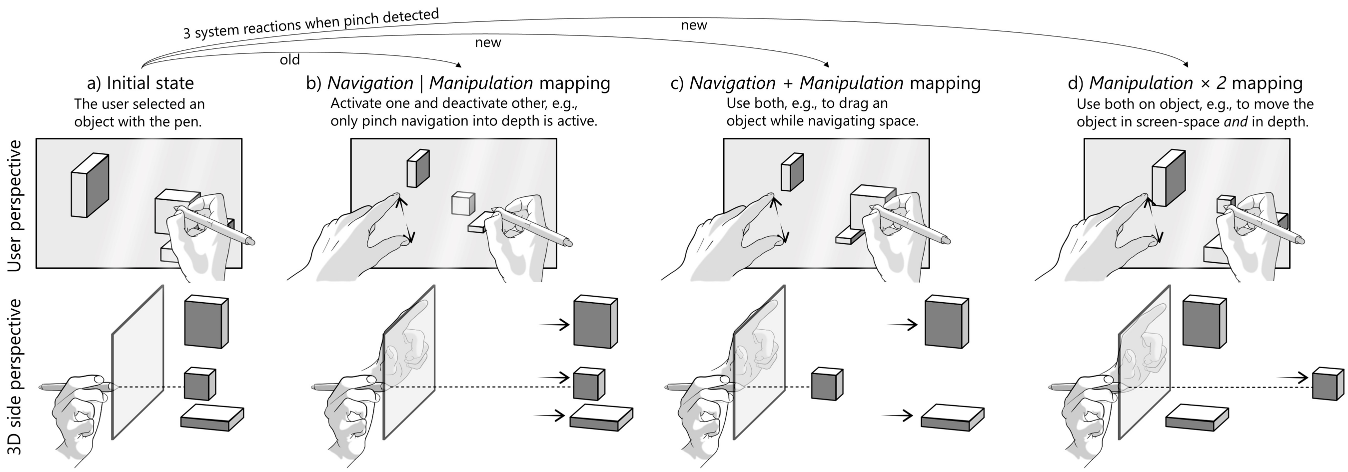

Figure 4: Navigation/Manipulation Mappings: The user begins with selecting a target (a). If the user engages a pinch gesture, typically only one mode is active (b). We propose two new mappings: users can drag an object via pen while navigating in 3D, which represents an intuitive extension from 2D to 3D (c); and users apply both pen and touch commands to one object for more expressive input (d).

The technique inherits from several prior concepts: (1) combining pen and touch to enable new tools [7, 22], (2) employ pinch-to-zoom to control object depth [20, 50], and (3) employ indirect touch to support precision and reduce occlusion [29, 38, 47]. This implies clutching for longer depth translations. This is relatively easy as the gesture is indirect and can be performed coarsely. For a clutch-free interaction, other gesture such as Cyclostar can be considered [36].

3.3 2D–3D Mapping Design Space

Simultaneous pinch-to-zoom (navigation) and pen dragging (manipulation) are complimentary activities. In parallel, a naive approach to directly employ the effect of both gestures can lead to a conflict. For example, the user selects a target, but then navigates to a view where the target is not visible. Should the target still be associated to the pen, or move with the navigation? There are several ways to implement the mappings in the context of a 3D scene, depending on what tasks and goals are important. To shed some light into this, Figure 4 illustrates a set of potential mappings that we detail in the following.

3.3.1 Navigation | Manipulation (Figure 4b). A unimodal mapping means only one of the tasks is supported and the other is not active. For example navigation only – i.e., after pen selection (a), a simultaneous pinch-to-zoom gesture deactivates pen input, so only one modality is active (b). This is useful when it is not desired to use both modalities in parallel, e.g., when no specific feature is supported for this. This mapping related to the initial state model that we described earlier (Figure 3a).

3.3.2 Navigation + Manipulation (Figure 4c). This mapping re-interprets the two-handed input to resolve the conflict. After an object is selected with the dominant hand (a), it remains associated to it even if the user navigates the canvas (c). This allows the user to hold or drag the object in screen-space while the user navigates the camera. Thus, useful when moving an object across large space. It is also applicable to sketching, as the line begins in the screen-space, concurrent navigation extends the line to 3D. A detailed state model is provided in Figure 3b, with the main difference being an mapping of a simultaneous touch gesture to 1DOF control of the pen-selected object.

3.3.3 Manipulation × 2 (Figure 4d). This mapping re-interprets the bi-manual input to apply to single object manipulation. Indirect touches of the non-dominant hand redirect toward the pen-selected object, for instance to enable 3DOF manipulation. The navigation that was associated to touch inputs, is not active for example. This is mainly useful if the user session involves frequent interaction with the object, e.g., to quickly drag & drop objects in 3D space. For a detailed state model, see Figure 3c. Different to the previous mapping, here simultaneous touches are directly mapped to RST tasks that affect the pen-selected object.

3.4 Visual Feedback

The touch gestures supporting the pen's actions can be implemented with various visual representations. Widgets make the task's operation more self-revealing and have been employed in several multi-DOF techniques [6, 55, 69]. But they also limit the user's input to a particular area on the screen. As an alternative, gestures such as pan & zoom are intuitive and apply to the entire UI, and do not necessitate a visual representation. As well, the concept of "indirect touch" can be efficient and relaxes the gesture placement [29, 38, 47]. Can simultaneous pen and touch be considered more as an intuitive gesture, or as a distinct feature that is better off with a widget? A 3DOF positioning task with two hands can lead to a high degree of compatibility between input and task, which may lower the need of a visual representation. When combining pen and touch for other 3DOF tasks, such as rotation-scale-translation tasks, graphical representations can become helpful to the user if not all transformations integrate.

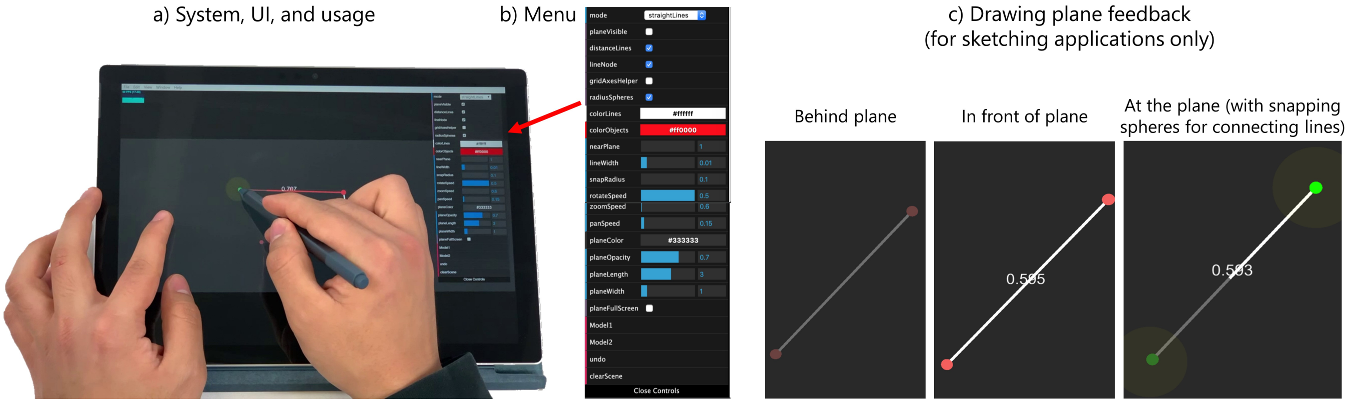

Figure 5: Our application prototype explores two-handed pen and touch tools (a). The menu allows to customise the GUI and select particular tools (b). Sketching happens on a transparent drawing plane in front of the screen-space, which is indicated through brightness feedback (c).

4 APPLICATIONS

We demonstrate applications of the pen and touch UI for 3D interactions. The applications explore interaction techniques along the following design dimensions:

- Regarding *Input States*, they explore the combination of the pen across touch based rotate, scale, and translate transformations.
- These are explored for two classes of *2D–3D Mappings*, firstly *Navigation + Manipulation* for 3D sketching and then *Manipulation x 2* for multi-DOF object manipulations.

The applications are explored in a prototype for sketching and object manipulation tasks across types of objects (i.e., a stroke, a graphical object, an object's face, or a vertex). We will first describe details on the application prototype, and then explore several interaction examples that are enabled by the system.

Apparatus. The prototype is is built on a 12.3" Microsoft Surface Pro 4 tablet in JavaScript using the three.js library, HTML and CSS. Figure 5 shows the application UI, menu, and visual feedback.

Menu. We utilise pen and touch inputs to cover the main interactions, and provide a menu for supportive features. An elegant UI can minimise menu usage as much as possible by using intuitive gestures, especially for a direct manipulation UI [22]. This avoid remembering where each mode is located in the menu, and its potential mode-switching cost. We use pen and touch for the core manipulation and navigation operations, and offer an additional menu to switch between some basic operations difficult with gestures. This provides a good balance between the simplicity/complexity of a UI, and its expressiveness in functionality. These include pen (stroke, fill color) and scene customisation (drawing plane's distance / color / opacity / toggle, feedback, grid, snapping size), and tool and object selection (whole object, object face only, a single vertex).

Feedback. 3D sketching on a drawing plane makes it particularly important for users to perceive the current 'slice' position in 3D space. There are several ways to address this, including showing a grid [64], and including depth cues, e.g., color, brightness, and saturation adjustments as objects are positioned further away. To aid 3D interaction, methods that ease the way users can translate and align objects in relation to the scene can be used, e.g., snapping mechanisms and magnetic guidelines [5]. As this is a general issue across 2D systems, we refer the interested reader to the prior art [42]. In our prototype, we use feedback to indicate the drawing plane position. It is see-through, making objects behind darker and objects in front brighter as usual. As an example, the line is grey if behind the drawing plane and white if in front of it. The user can configure the transparency level of the plane and toggle it on or off.

4.1 *Navigation + Manipulation* for 3D Sketching

The following examples are based on the *Navigation + Manipulation*, i.e., they empower the user to combine both tasks for integrated operations. This is demonstrated on the example of sketching lines and extending those to 3D space. Figure 6 (a) to (d) show screenshots of the use and results of the prototype application.

4.1.1 Freelines + Zoom (Figure 6a). We demonstrate a mode where users simply draw free lines with the pen and extend them to 3D. The dominant hand with the pen initiates drawing the 2D circle (see Figure 6b (1)). When issuing pinch-to-zoom, the user navigates the canvas – this, in parallel to drawing a line, will extend it to the third dimension (2). It feels like a new experience of drawing, as one gains the possibility to rapidly draw in 3D, yet with a different spatial frequency and precision for the Z axis (2DOF dominant hand vs. 1DOF nondominant hand with clutching). In principle, pinch-to-zoom provides high spatial precision, but of course it is different to simply drawing a 2D line, needing coordination of both hands and cognitively unifying the inputs to a 3D sketch. We currently see the main application in quick & dirty sketching of complex 3D strokes, as new forms of 3D lines – for instance a spiral line (2) – become possible, as the user flexibly adapts line generation during sketching toward any 3D direction.

4.1.2 FreeLines + Rotate (Figure 6b). What happens when rotating while drawing? Holding the pen at one position while rotating in 2D allows to draw a perfect circle, c.f. Figure 6b (1). A different rotation direction in 3D space extends the circle toward the depth of the world. In succession with different directions, users can outline the round mesh of a sphere (3); something that, on a typical sketch UI, is difficult to achieve in a short time.

4.1.3 FreeLines + Panning. This combination makes it possible for the user to draw only by panning with touch. This is a straightforward extension of a sketched line; panning will change the underlying 2D plane and extend the line to a new position (if the pen is not moved exactly parallel to the finger). The advantage is that users can extend the line beyond the screen's boundaries, potentially infinitely across the canvas, and also offload the physical effort of the dominant hand, to indirect touch dragging anywhere on the UI.

4.1.4 StraightLines + Zoom (Figure 6c). Via constraints, we show how users can precisely create lines and connect them. We focus on straight lines as the pen is on the screen. A snapping mechanism allows users to easily connect a line to a vertex of another line (or, other objects with vertices). Figure 6c shows a single line example, where users connect one vertex to another in 3D. The user starts the line by selecting an existing vertex via pen (1). For each line, the depth level feedback includes a label in the center that specifies distance to the drawing plane. The user then zooms out (2), until the feedback shows that the canvas is at the same level as the destination vertex by the green colored dot (3). Then, the user connects the line to this vertex to finish. Thus, one can consider this an alternative technique to depth based sketching, where sketches project directly to a 3D surface. It lends itself more to drawing a line freely into blank space. It affords composition of lines to create 3D shapes such as cake slices (4). Note that in freeform sketching, no constraints are involved. When pinching, the line will be orthogonal to the view plane as the pen is does not move.

4.1.5 Outline + Zoom (Figure 6d). This tool demonstrates the creation of a 3D object with a single stroke. In 2D applications, often users can create a rectangle by drawing its diagonal. Here we utilise the simultaneous, alternating use of the modalities to extend this to 3D space. The user draws a 2D rectangle on the drawing plane, by defining its diagonal as shown in Figure 6d (1) to (2). Without lifting the stylus, the user can touch down two fingers to initiate pinch-to-zoom. The change in depth is automatically translated to the currently drawn rectangle, directly extruding the rectangle face by the level of zooming (3). In addition, users can now move the pen in the two dimensions on the screen to change the end position of the extrusion. As a result, the user can rapidly create a cuboid or parallelogram (4), and the principle idea to create 3D objects by strokes is likely to extend to other object types, too. Note that we use a mode switch to activate this mode, but it could also be a one-time-mode like creating a shape in Microsoft Powerpoint.

4.2 *Manipulation x 2* for Object Manipulation

The following examples are based on the *Manipulation × 2* – i.e., they take advantage of focusing both hands for a more powerful manipulation of a single object. This is demonstrated based on manipulations on objects and their shape for translation, rotation, and scaling transformations. Figure 6 (e) to (g) show examples and outcomes from the implemented prototype.

Notably, the rotation and scaling features of (f, g) are supported by widgets, allowing single-touch to be used for performing the actions. We opted for a basic design (yellow area), which is not self-intuitive at this stage but useful for the interaction. The design can be improved further by applying prior work on this case [49].

4.2.1 3D Drag & Drop (Figure 6e). This example presents a more integrated experience where both hands operate on the same object in the same task – drag & drop in 3D. The initial selection is done by the pen – the object is selected in the 3D space by a ray cast from the pen position.

On the example of moving a piece of a Rubik's cube, we demonstrate how the technique affords an easy method to 3D translate (Figure 6e). The user initially draws a piece outwards, towards the front and left side of the space by a pinch-in gesture with slight pen dragging (1). This is followed by horizontal pen dragging (2), as the user inspects the drop destination (the middle cell of the Rubik's cube's right side). A pinch-in gesture moves the target toward it (3), and combined with a left-drag of the pen, the user can precisely position the piece into the larger model (4).

The task itself involved many sub-steps: translation in a variety of directions as well as pauses in between to cognitively process the movement and plan the next steps. The technique integrates those intuitively to fully control the object's movement between 2D and 3D spaces, without a need to release the tip from the screen.

4.2.2 Extrude + Rotate (Figure 6f). Extrusion modelling often begins with a basic shape as a point of departure, from which the user extrudes and transforms in creative ways. This technique demonstrates a way to integrate rotation of an object at the same time when performing an extrusion.

Figure 6f describes a single process where one of the faces of an existing model is extended. Face selection is established by direct pen contact (a). Extrusion is, only for pen dragging, enabled in screen-space. As the user wants to move it a bit forward, the user can use thumb dragging, to extrude in the depth dimension toward the camera (b). By using the secondary widget, the user rotates the pen-selected face (c). The widget is designed to activate when the finger starts within the area, and the direction of finger dragging is mapped to rotate the face (2DOF rotation). The user can arbitrarily aggregate these actions to complete a larger model (d).

4.2.3 Extrude + Scale (Figure 6g). Scale transformations complement rotation and translation. In principle, all of such tasks can be used in a dynamic interplay. This example shows how scaling integrates in an example workflow of creating a custom table design. The example is inspired by those presented in MockupBuilder [14].

Figure 6g presents the steps to create such a shape. First, the user selects the model's face (1) and extrudes it (2) in two dimensional space with a pen. Immediately, the user can scale the newly extruded face to inspect several table surface sizes. As translation and rotation is available through the two widgets, the scaling is enabled through indirect touch dragging on the canvas. The shape is scaled uniformly, although in principle the system could further distinguish the dragging direction to scale width/height. After the operation, the user found the right size for their custom table (4).

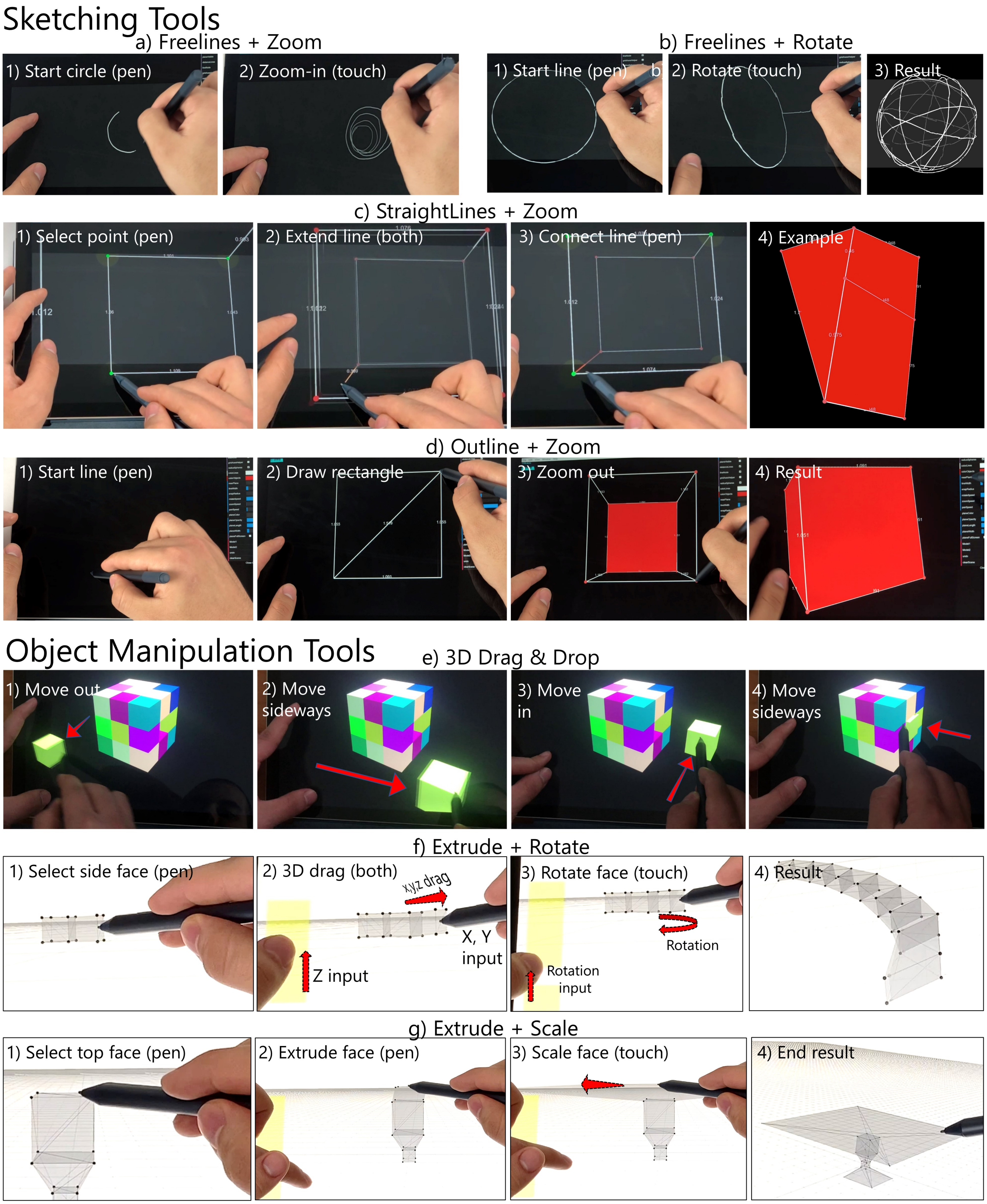

Figure 6: Techniques implemented in our application, empowering the user in the manipulation of 3D content. The top half presents applications for sketching (a-d), and the bottom half for 3D object transformations (e-g).

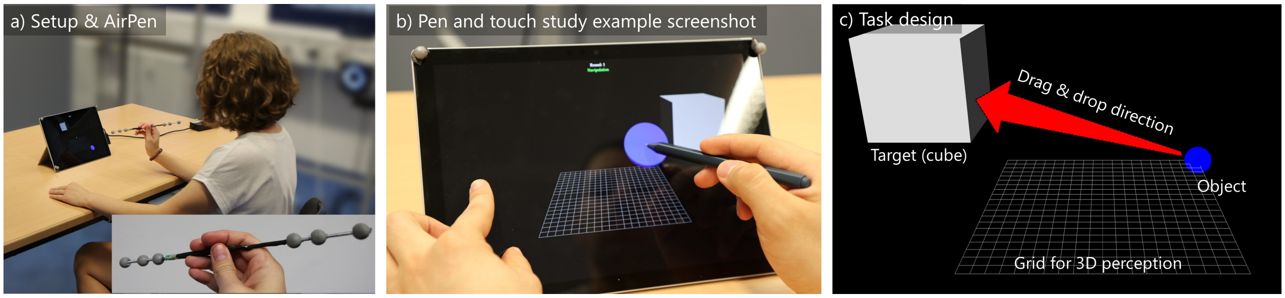

Figure 7: Our user study's setup (a), an example of interacting via pen and touch condition (b), and task illustration (c).

5 USER STUDY

The applications we developed showed the qualitative strengths of pen and touch interaction for 3D tools. As complement, we conducted an empirical user study to assess the core task of 3D dragging, a task that is integral to many of the proposed applications. In particular, it is closely related to translation tasks from one to another endpoint (Drag & Drop, StraightLines + Zoom, Outline + Zoom), and partially related to Extrude + Rotate / Scale where users coordinate multiple DOFs to achieve a final RST value. The goal is to compare pen and touch to a clear DOF separation approach (widgets) and a DOF integration approach (3D mid-air dragging).

5.1 Task

The study task was adapted from Balakrishnan and Kurtenbach's object translation study of desktop devices [3]. This is because their research question is similar, yet in a different context (mouse-based input). We opted to keep the task goal the same: users cannot perform errors and the task only finishes as docking is complete.

The basic task is to move a sphere from one corner to the opposing diagonal corner. In that corner, the user fits the sphere into a target cube (Figure 7c). A flat grid is shown in the background for 3D perception. There are 8 starting points (from each corner) that are randomly assigned to the 8 trials. The object is a round sphere and the target is a transparent cube. The cube's width is 11.67 world units and the sphere's radius is about 7 world units. The task finishes when the sphere is completely within the cube and users release the pen or key (in the mid-air condition). Visual feedback is given to indicate the state. The sphere is blue normally, yellow when overlapping with the cube, and green when inside. The camera is kept still, only the object can be moved.

5.2 Techniques

Our study involves three techniques.

5.2.1 PEN-AND-TOUCH. The first technique is the Bi-3D 3DOF translation technique as described before (see Figure 6e). An example of the study is shown in Figure 7b.

5.2.2 WIDGET. This method is based on a clear DOF separation, similar to handle boxes [23] or Arcball [53]. The widget provides handles for each DOF, and also for each 2DOF translation (see Figure 8). To conduct a 3D translation, the user can either independently manipulate each dimension, or combine with 2DOF tools. It is oriented toward the camera so that users can translate all axis in relation to screen-space.

5.2.3 AIR-PEN. This technique allows users to drag & drop an object by spatial 3D movement, as illustrated by figure 9. The design of this baseline has been informed through prior work. Mid-air input has often been used to extend desktop systems [52, 57], interactive surfaces [4, 60], and 3D CAD [14, 27, 43]. Here, raypointing selection is a suitable candidate, where users direct the handheld device to the desired screen position [11, 45, 46]. The *AIR-PEN* uses this selection, to then perform dragging. In particular, selection happens by 1) pointing the pen in the dominant hand in mid-air at the target on the screen, 2) pressing the space button on a separate keyboard with the non-dominant hand to confirm. Then, spatial 3D movement of the pen will 1:1 translate to movement of the object in the 3D scene. Meaning, horizontal and vertical movement translates the object parallel to screen-space, and movement toward/away from the screen manipulates object depth.

To track the 3D movement and direction of the pen, we used an Optitrack motion tracking system with 8 cameras around the table. A first prototype used markers on the stylus for tracking, but this often led to tracking loss as the user was handling the pen. We therefore replaced the pen with a metallic rod long enough to support several markers robustly. The center part was taped to provide an easier grip, but it did not introduce a noticeable weight difference to the pen.

5.3 Study Design

The study uses a repeated measure within-subject design. The main independent variable was technique, including *WIDGET*, *PEN-AND-TOUCH*, and *AIR-PEN*. The order of the techniques was counterbalanced using a Latin square. Each technique session includes 24 trials (8 diagonal directions × 3 repetitions), resulting in 72 tasks per user. Users could take a break anytime and continue. Mostly users finished 24 trials in one go.

5.4 Apparatus

We use the same system as described in the applications.

5.5 Procedure

Participants were introduced to the scope of the project. Then, they were asked to sign a consent form and fill in a demographics questionnaire. Participants were given a few trials as training before each technique until they felt comfortable to proceed with the

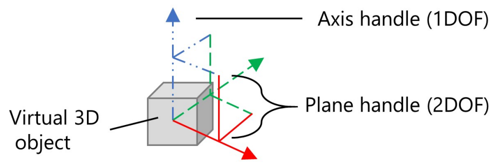

Figure 8: *Widget*: The first baseline technique allows object translation through DOF separation. Each handle allows 1 or 2D axis control.

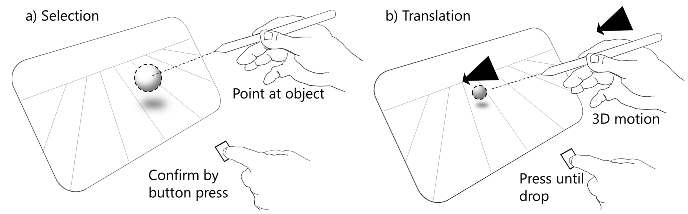

Figure 9: *Air-Pen*: The second baseline technique allows 3D drag & drop by a 3D gesture. The user selects the target by mid-air pointing and pressing a button (a), and then spatial 3D motion of the pen translates the object (b).

actual experiment (2–4 trials for most users). After each technique, a usability questionnaire was filled out. At the end, participants were asked to rank all the techniques and discuss their reasons. The study duration took on average 30 minutes for each session.

5.6 Evaluation Methods

- **Task Completion Time**: Measures the user's performance. It is the time taken by the participant from the start of the trial (object and target appears), until the task finishes (target is inside the cube and input is released).
- **Integrality Ratio**: To understand the DOF integration, we assess the integrality/separability ratio as defined by Jacob [25]. For this, the movement trajectories of the dragged object are segmented into equal time (16.6ms in our case). Each segment is classified to either Euclidian (if movement occurred in multiple dimensions) or city-block (if movement occurred in one dimension). The degree of integrality is computed by the ratio of Euclidean to city-block movements. Values greater than 1 indicate higher Euclidean (integral) than city-block (separable) results.
- **Usability Ratings**: Users rate the usability of the techniques for 6 statements on a 5-Point Likert scale from 1 (strongly disagree) to 5 (strongly agree). The 6 statements were: *"Performing the task with this technique was [easy to use | easy to learn | without physical effort | precise | fast | useful for 3D manipulation"].*
- **Feedback**: Users comment in the questionnaire and discussed the techniques in a short interview.

5.7 Participants

We recruited 18 paid participants (9 female) with an average age of 25.27 years (SD = 3.22) using university mailing lists. Participants had technical backgrounds (IT, Engineering, computer science). Few users reported using a digital pen regularly. However many have experience with 3D modelling through university courses, hobby, or job. The users were also experienced with using tablets.

5.8 Results

Over most of the factors, the results indicate that in this particular task, the *Widget* approach is less preferable than the bi-manual and mid-air technique (wrt. task completion time, integrality, user ratings and comments). However, there were surprisingly few statistically significant differences between *Pen-and-Touch* and *Air-Pen*. Why is that?

5.8.1 Task Completion Time. A repeated measures ANOVA with Greenhouse-Geisser correction revealed a significant main effect of the interaction technique on task duration ($F_{2,34} = 21.79$, $p < .001$). Post-hoc comparisons with Bonferroni correction show *Widget* (M = 22.25; SD = 14.53) required significantly more time to complete the task than *Pen-and-Touch* (M = 7.87; SD = 3.65) and *Air-Pen* (M = 8.93; SD = 3.43, both p < .001). No statistical significance was found between *Pen-and-Touch* and *Air-Pen* (p = .63).

5.8.2 Integrality Analysis. A repeated measures ANOVA showed statistical significant effects of the technique on the integrality ratio ($F_{2,30}$ = 26.382; p < .001). Pairwise comparison with Bonferroni correction shows statistical differences (p < .05) between the *Widget* (M=.41; SD=.42) and *Air-Pen* (M=1.75; SD=.42), and between *Widget* and *Pen-and-Touch* (M=1.44; SD=.65). This statistically shows that both *Pen-and-Touch* and *Air-Pen* resulted in a significantly higher degree of integrality than *Widget*.

5.8.3 User Ratings. A non-parametric Friedman test and post-hoc Wilcoxon signed rank test with Bonferroni corrections were used for analysing the ratings. No significant main effect was found between the three modalities for physical effort ($\chi^2(19) = 1.97$, p = .37). We found significant main differences for perceived speed ($\chi^2(19) = 18.97$, p < .001), ease of use ($\chi^2(19) = 17.71$, p < .001), learnability ($\chi^2(19) = 18.82$, p < .001), precision, ($\chi^2(19) = 17.32$, p < .001), and usefulness for 3D manipulation ($\chi^2(19) = 14.46$, p = .001). In those categories, users rated *Air-Pen* and *Pen-and-Touch* significantly higher than the baseline (all p < .05). No significant differences were found between *Air-Pen* and *Pen-and-Touch*.

5.8.4 User Feedback. Many users provided comments for all the techniques, mostly confirming the trend from the quantitative analysis. The participants' most important factors were ease of use and comfort. For *Pen-and-Touch* and *Air-Pen* user stated comments like P3: *"it is easy and comfortable to perform"*. However, as a potential disadvantage, several users pointed out physical demand of *Air-Pen* (P8: *"it is the hardest one physically"*). Regarding precision, users' perceptions were mixed, as on one hand some found *Pen-and-Touch* less precise than *Air-Pen*, but on the other hand there were users who found the interaction technique *"surprisingly intuitive and precise"*. For the interaction with the *Widget* technique, a participant reported that they use the editor in Unity quite often with the mouse, and therefore found it easy to use as it is the same widget. Of all, two users also mentioned it is very precise, which hints to the benefits of DOF separation. Users with less experience had issues to get used to the technique: *"'I'd probably need to practice a little more to get used to it"* (P15).

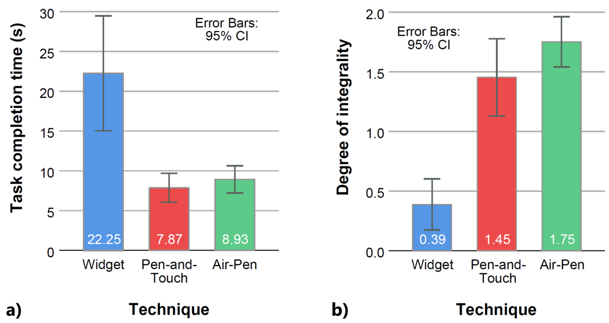

Figure 10: User study results on task completion time (a) and degree of integrality (b).

6 DISCUSSION

We investigated bi-manual pen and touch interaction for the manipulation of 3D objects. We examined design considerations, applications, and empirical results. Overall, our work reveals that there are yet many opportunities to design for a more seamless interplay between the two hands. 3D interaction can benefit from asymmetric bi-manual input modes and pen interaction can benefit from extension to the third dimension.

Our experimental comparison provided first insights into the user performance with such Bi-3D interaction, by taking a look at a 3D translation task. This task was designed to be a typical program experience on a computer with the main application window showing the virtual 3D world through the screen. Results were interesting, as first of all showing how the bimanual UI can be used in this instance as a direct way to drag & drop a virtual object faster than when using the widgets. As well, that statistically there was no significant difference between the mid-air and two handed approach. These are promising results, that really make us believe whether these techniques can be come a practical addition to tablet devices. Qualitative feedback by the participants was positive, with users preferring the technique for its easy and intuitive way of use.

To explore use cases for Bi-3D, we build our application that showcased several interesting tool applications that particularly nurture from exploiting the dynamic interplay between the hands. While we refer to the corresponding section to better assess each techniques potential strength and weakness for the proposed task context, we want to highlight the two classes of potential applications. First, the more typical canvas and drawing tools where pinch-to-zoom provides a simple but effective addition that aligns with what users expect of their combined interaction. This is a simplified but also a novel way of interaction, as it directly integrates in the existing workflow and extends it to with a third dimension. Second, for rather expert use, to consider their combination in object manipulations like extrusion, rotate, scale and translate actions. Here the proposed tools can be used complementary, as an additional UI mode or integrated in existing widgets.

There are several points that we want to further discuss, as limitations of the current and opportunities for future work.

- Evaluation: We focused on controlled evaluation of 3D translation, but the concrete application prototype is broader and still needs to be tested in practical use. As commercial applications exist, it would be interesting to compare whole tools. Other baselines are as well interesting, e.g., many applications use multiple viewports that decompose the 3D task into 2D views, which alleviates some of the widget-only issues. Another interesting baseline for empirical comparison includes purely touch based 3DOF techniques [39, 50].
- Bimanual Coordination: Coordination can be relatively high for some techniques, but it's also key to render complex 3D tasks easier to perform. Our interaction is principally designed to reduce demand, compared to e.g. widgets or 3D-touch techniques, as pinch-to-zoom and the pen combine to 3D input, but full simultaneous use may be more difficult. The benefit is that users can anytime easily shift between manipulating dimensions in parallel or in sequence (opposed to, e.g., widgets are exclusively designed for sequential use).
- Precision and Coordination: Our evaluation focused on docking tasks, but without answering fundamental trade-offs in precision and coordination demands across the range of tasks from precise freeform 3D sketching to relatively simple object docking. This is important to understand the level of parameters and constraints on the DOF and the two hands' input for this category of Bi-3D techniques in practice.
- UI Modes: Our application prototype uses modes accessible through the menu. This introduces a cost to remember where the modes are and to switch between them. Bi-3D includes two coherent gesture sets that cover core navigation and manipulation tasks, but parameters like object type and UI options are available by the menu.
- Technique Design: How can the techniques become more self-revealing, to include a broader set of user groups? Part of the interactions take advantage of what users may expect from the combined gestures (e.g., that a simultaneous pinch-to-zoom will extend a 2D operation to 3D). Prior work suggested several designs that are potential candidates to provide support, e.g., a context aware widget that adapts to the hand's posture and location on the screen.

7 CONCLUSION

The increasing advances of multi-modal interfaces for tablets and other interactive surfaces gives rise to new possibilities that empower the user in their interaction with digital content. 3D manipulation is an intricate topic with the challenge to coordinate multiple degrees-of-freedom with the control of multiple dimensions of virtual objects. This paper proposed pen and touch based interaction techniques for manipulation of 3D objects. We contribute screen-space formulations, interaction techniques, applications, and empirical data on user performance as an extensive exploration of this space. We hope that with our work, we can raise the utility and suitability of pen and touch interactions for designing 3D content, and open more possibilities of design beyond the desktop toward the broad range of usage contexts that tablets afford.

8 ACKNOWLEDGEMENTS

The work was supported by the German Research Foundation (DFG) under project no. 425869382 and by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [Voice of Wisdom].

REFERENCES

- [1] Seok-Hyung Bae, Ravin Balakrishnan, and Karan Singh. 2008. ILoveSketch: As-natural-as-possible Sketching System for Creating 3D Curve Models. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (Monterey, CA, USA) *(UIST '08)*. ACM, New York, NY, USA, 151–160. https://doi.org/10.1145/1449715.1449740
- [2] Ravin Balakrishnan, George Fitzmaurice, Gordon Kurtenbach, and William Buxton. 1999. Digital Tape Drawing. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology* (Asheville, North Carolina, USA) *(UIST '99)*. Association for Computing Machinery, New York, NY, USA, 161–169. https://doi.org/10.1145/320719.322598
- [3] Ravin Balakrishnan and Gordon Kurtenbach. 1999. Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. ACM, New York, NY, USA, 56–62. https://doi.org/10.1145/302979.302991
- [4] Amartya Banerjee, Jesse Burstyn, Audrey Girouard, and Roel Vertegaal. 2011. Pointable: An In-air Pointing Technique to Manipulate Out-of-reach Targets on Tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Kobe, Japan) *(ITS '11)*. ACM, New York, NY, USA, 11–20. https://doi.org/10.1145/2076354.2076357
- [5] Michel Beaudouin-Lafon and Henry Michael Lassen. 2000. The Architecture and Implementation of CPN2000, a Post-WIMP Graphical Application. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology* (San Diego, California, USA) *(UIST '00)*. Association for Computing Machinery, New York, NY, USA, 181–190. https://doi.org/10.1145/354401.354761
- [6] Hrvoje Benko and Steven Feiner. 2007. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. In *2007 IEEE symposium on 3D user interfaces (3D UI '09)*. IEEE. https://doi.org/10.1109/3DUI.2007.340778
- [7] Peter Brandl, Clifton Forlines, Daniel Wigdor, Michael Haller, and Chia Shen. 2008. Combining and Measuring the Benefits of Bimanual Pen and Direct-touch Interaction on Horizontal Interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (Napoli, Italy) *(AVI '08)*. ACM, New York, NY, USA, 154–161. https://doi.org/10.1145/1385569.1385595
- [8] Thomas Butkiewicz, Andrew H. Stevens, and Colin Ware. 2019. Multi-Touch 3D Positioning with the Pantograph Technique. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (Montreal, Quebec, Canada) *(I3D '19)*. Association for Computing Machinery, New York, NY, USA, Article 13, 9 pages. https://doi.org/10.1145/3306131.3317014
- [9] W. Buxton and B. Myers. 1986. A Study in Two-handed Input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, Massachusetts, USA) *(CHI '86)*. ACM, New York, NY, USA, 321–326. https://doi.org/10.1145/22627.22390
- [10] Drini Cami, Fabrice Matulic, Richard G. Calland, Brian Vogel, and Daniel Vogel. 2018. Unimanual Pen+Touch Input Using Variations of Precision Grip Postures. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 825–837. https://doi.org/10.1145/3242587.3242652
- [11] Xiang Cao and Ravin Balakrishnan. 2003. VisionWand: Interaction Techniques for Large Displays Using a Passive Wand Tracked in 3D. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (Vancouver, Canada) *(UIST '03)*. Association for Computing Machinery, New York, NY, USA, 173–182. https://doi.org/10.1145/964696.964716
- [12] Jonathan M. Cohen, Lee Markosian, Robert C. Zeleznik, John F. Hughes, and Ronen Barzel. 1999. An Interface for Sketching 3D Curves. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics* (Atlanta, Georgia, USA) *(I3D '99)*. Association for Computing Machinery, New York, NY, USA, 17–21. https://doi.org/10.1145/300523.300655
- [13] Philip L. Davidson and Jefferson Y. Han. 2008. Extending 2D Object Arrangement with Pressure-Sensitive Layering Cues. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (Monterey, CA, USA) *(UIST '08)*. Association for Computing Machinery, New York, NY, USA, 87–90. https://doi.org/10.1145/1449715.1449730
- [14] Bruno R. De Araùjo, Géry Casiez, and Joaquim A. Jorge. 2012. Mockup Builder: Direct 3D Modeling on and above the Surface in a Continuous Interaction Space. In *Proceedings of Graphics Interface 2012* (Toronto, Ontario, Canada) *(GI '12)*. Canadian Information Processing Society, CAN, 173–180.
- [15] Julie Dorsey, Songhua Xu, Gabe Smedresman, Holly Rushmeier, and Leonard McMillan. 2007. The Mental Canvas: A Tool for Conceptual Architectural Design and Analysis. In *Proceedings of the 15th Pacific Conference on Computer Graphics and Applications (PG '07)*. IEEE Computer Society, Washington, DC, USA, 201–210. https://doi.org/10.1109/PG.2007.62
- [16] Tomas Dorta, Gokce Kinayoglu, and Michael Hoffmann. 2014. Hyve-3D: A New Embodied Interface for Immersive Collaborative 3D Sketching. In *ACM SIGGRAPH 2014 Studio* (Vancouver, Canada) *(SIGGRAPH '14)*. Association for Computing Machinery, New York, NY, USA, Article 37, 1 pages. https://doi.org/10.1145/2619195.2656325
- [17] Yotam Gingold, Takeo Igarashi, and Denis Zorin. 2009. Structured Annotations for 2D-to-3D Modeling. *ACM Trans. Graph.* 28, 5 (Dec. 2009), 1–9. https://doi.org/10.1145/1618452.1618494
- [18] Yves Guiard. 1987. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. In *Journal of Motor Behavior*, Vol. 19. 486–517.
- [19] Mark Hancock, Sheelagh Carpendale, and Andy Cockburn. 2007. Shallow-Depth 3d Interaction: Design and Evaluation of One-, Two- and Three-Touch Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 1147–1156. https://doi.org/10.1145/1240624.1240798
- [20] Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. 2009. Sticky Tools: Full 6DOF Force-Based Interaction for Multi-Touch Tables. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Banff, Alberta, Canada) *(ITS '09)*. Association for Computing Machinery, New York, NY, USA, 133–140. https://doi.org/10.1145/1731903.1731930
- [21] Rorik Henrikson, Bruno De Araujo, Fanny Chevalier, Karan Singh, and Ravin Balakrishnan. 2016. Storeoboard: Sketching Stereoscopic Storyboards. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4587–4598. https://doi.org/10.1145/2858036.2858079
- [22] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko, and Bill Buxton. 2010. Pen + Touch = New Tools. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. ACM, New York, NY, USA, 27–36. https://doi.org/10.1145/1866029.1866036
- [23] Stephanie Houde. 1992. Iterative Design of an Interface for Easy 3-D Direct Manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Monterey, California, USA) *(CHI '92)*. Association for Computing Machinery, New York, NY, USA, 135–142. https://doi.org/10.1145/142750.142772
- [24] Takeo Igarashi, Satoshi Matsuoka, and Hidehiko Tanaka. 1999. Teddy: A Sketching Interface for 3D Freeform Design. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '99)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 409–416. https://doi.org/10.1145/311535.311602
- [25] Robert J. K. Jacob, Linda E. Sibert, Daniel C. McFarlane, and M. Preston Mullen, Jr. 1994. Integrality and Separability of Input Devices. *ACM Trans. Comput.-Hum. Interact.* 1, 1 (March 1994), 3–26. https://doi.org/10.1145/174630.174631
- [26] Kiia Kallio. 2005. 3D6B Editor: Projective 3D Sketching with Line-Based Rendering. In *Eurographics Workshop on Sketch-Based Interfaces and Modeling*, Joaquim Armando Pires Jorge and Takeo Igarashi (Eds.). The Eurographics Association. https://doi.org/10.2312/SBM/SBM05/073-079
- [27] Yongkwan Kim, Sang-Gyun An, Joon Hyub Lee, and Seok-Hyung Bae. 2018. Agile 3D Sketching with Air Scaffolding. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 238, 12 pages. https://doi.org/10.1145/3173574.3173812
- [28] Kenrick Kin, Maneesh Agrawala, and Tony DeRose. 2009. Determining the Benefits of Direct-Touch, Bimanual, and Multifinger Input on a Multitouch Workstation. In *Proceedings of Graphics Interface 2009* (Kelowna, British Columbia, Canada) *(GI '09)*. Canadian Information Processing Society, CAN, 119–124.
- [29] Sebastian Knoedel and Martin Hachet. 2011. Multi-Touch RST in 2D and 3D Spaces: Studying the Impact of Directness on User Performance. In *Proceedings of the 2011 IEEE Symposium on 3D User Interfaces (3DUI '11)*. IEEE Computer Society, USA, 75–78.
- [30] P. Lapides, E. Sharlin, M. C. Sousa, and L. Streit. 2006. The 3D Tractus: a three-dimensional drawing board. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)*. 8 pp.–.
- [31] Celine Latulipe, Craig S. Kaplan, and Charles L. A. Clarke. 2005. Bimanual and Unimanual Image Alignment: An Evaluation of Mouse-Based Techniques. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (Seattle, WA, USA) *(UIST '05)*. Association for Computing Machinery, New York, NY, USA, 123–131. https://doi.org/10.1145/1095034.1095057
- [32] Celine Latulipe, Stephen Mann, Craig S. Kaplan, and Charlie L. A. Clarke. 2006. SymSpline: Symmetric Two-Handed Spline Manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 349–358. https://doi.org/10.1145/1124772.1124825
- [33] Jinha Lee and Hiroshi Ishii. 2010. Beyond: Collapsible Tools and Gestures for Computational Design. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. Association for Computing Machinery, New York, NY, USA, 3931–3936. https://doi.org/10.1145/1753846.1754081
- [34] Jingbo Liu, Oscar Kin-Chung Au, Hongbo Fu, and Chiew-Lan Tai. 2012. Two-Finger Gestures for 6DOF Manipulation of 3D Objects. *Comput. Graph. Forum* 31, 7pt1 (Sept. 2012), 2047–2055. https://doi.org/10.1111/j.1467-8659.2012.03197.x
- [35] Pedro Lopes, Daniel Mendes, Bruno Araújo, and Joaquim A. Jorge. 2011. Combining Bimanual Manipulation and Pen-based Input for 3D Modelling. In *Proceedings of the Eighth Eurographics Symposium on Sketch-Based Interfaces and Modeling* (Vancouver, British Columbia, Canada) *(SBIM '11)*. ACM, New York, NY, USA,

15–22. https://doi.org/10.1145/2021164.2021168

[36] Sylvain Malacria, Eric Lecolinet, and Yves Guiard. 2010. Clutch-Free Panning and Integrated Pan-Zoom Control on Touch-Sensitive Surfaces: The Cyclostar Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2615–2624. https://doi.org/10.1145/1753326.1753724

[37] Anthony Martinet, Gery Casiez, and Laurent Grisoni. 2010. The design and evaluation of 3d positioning techniques for multi-touch displays. In *2010 IEEE symposium on 3D user interfaces (3DUI)*. IEEE, 115–118.

[38] Anthony Martinet, Géry Casiez, and Laurent Grisoni. 2010. The Effect of DOF Separation in 3D Manipulation Tasks with Multi-Touch Displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (Hong Kong) *(VRST '10)*. Association for Computing Machinery, New York, NY, USA, 111–118. https://doi.org/10.1145/1889863.1889888

[39] Anthony Martinet, Gery Casiez, and Laurent Grisoni. 2012. Integrality and Separability of Multitouch Interaction Techniques in 3D Manipulation Tasks. *IEEE Transactions on Visualization and Computer Graphics* 18, 3 (March 2012), 369–380. https://doi.org/10.1109/TVCG.2011.129

[40] Fabrice Matulic and Moira Norrie. 2012. Empirical Evaluation of Uni- and Bimodal Pen and Touch Interaction Properties on Digital Tabletops. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. ACM, New York, NY, USA, 143–152. https://doi.org/10.1145/2396636.2396659

[41] Fabrice Matulic and Moira C. Norrie. 2013. Pen and Touch Gestural Environment for Document Editing on Interactive Tabletops. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces* (St. Andrews, Scotland, United Kingdom) *(ITS '13)*. ACM, New York, NY, USA, 41–50. https://doi.org/10.1145/2512349.2512802

[42] Daniel Mendes, Fabio Marco Caputo, Andrea Giachetti, Alfredo Ferreira, and J Jorge. 2019. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, Vol. 38. Wiley Online Library, 21–45.

[43] D. Mendes, F. Fonseca, B. Araùjo, A. Ferreira, and J. Jorge. 2014. Mid-air interactions above stereoscopic interactive tables. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 3–10. https://doi.org/10.1109/3DUI.2014.6798833

[44] M. Nebeling and M. Speicher. 2018. The Trouble with Augmented Reality/Virtual Reality Authoring Tools. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 333–337. https://doi.org/10.1109/ISMAR-Adjunct.2018.00098

[45] Dan R. Olsen and Travis Nielsen. 2001. Laser Pointer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '01)*. ACM, New York, NY, USA, 17–22. https://doi.org/10.1145/365024.365030

[46] J. Karen Parker, Regan L. Mandryk, and Kori M. Inkpen. 2005. TractorBeam: Seamless Integration of Local and Remote Pointing for Tabletop Displays. In *Proceedings of Graphics Interface 2005* (Victoria, British Columbia) *(GI '05)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 33–40.

[47] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-Touch: Combining Gaze with Multi-Touch for Interaction on the Same Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 509–518. https://doi.org/10.1145/2642918.2647397

[48] Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2016. Partially-indirect Bimanual Input with Gaze, Pen, and Touch for Pan, Zoom, and Ink Interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 2845–2856. https://doi.org/10.1145/2858036.2858201

[49] Ken Pfeuffer, Ken Hinckley, Michel Pahud, and Bill Buxton. 2017. Thumb + Pen Interaction on Tablets. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 3254–3266. https://doi.org/10.1145/3025453.3025567

[50] Jason L. Reisman, Philip L. Davidson, and Jefferson Y. Han. 2009. A Screen-Space Formulation for 2D and 3D Direct Manipulation. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) *(UIST '09)*. Association for Computing Machinery, New York, NY, USA, 69–78. https://doi.org/10.1145/1622176.1622190

[51] Hugo Romat, Emmanuel Pietriga, Nathalie Henry-Riche, Ken Hinckley, and Caroline Appert. 2019. SpaceInk: Making Space for In-Context Annotations. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 871–882. https://doi.org/10.1145/3332165.3347934

[52] Emanuel Sachs, Andrew Roberts, and David Stoops. 1991. 3-Draw: A Tool for Designing 3D Shapes. *IEEE Comput. Graph. Appl.* 11, 6 (Nov. 1991), 18–26. https://doi.org/10.1109/38.103389

[53] Ken Shoemake. 1992. ARCBALL: A User Interface for Specifying Three-Dimensional Orientation Using a Mouse. In *Proceedings of the Conference on Graphics Interface '92* (Vancouver, British Columbia, Canada). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 151–156. https://doi.org/10.5555/155294.155312

[54] Arjun Srinivasan, Bongshin Lee, Nathalie Henry Riche, Steven M. Drucker, and Ken Hinckley. 2020. InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376782

[55] Sven Strothoff, Dimitar Valkov, and Klaus Hinrichs. 2011. Triangle Cursor: Interactions with Objects above the Tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Kobe, Japan) *(ITS '11)*. Association for Computing Machinery, New York, NY, USA, 111–119. https://doi.org/10.1145/2076354.2076377

[56] Mental Canvas Authoring System. Accessed 7 April, 2021. Mental Canvas Authoring System. http://www.mentalcanvas.com.

[57] Robert Wang, Sylvain Paris, and Jovan Popović. 2011. 6D Hands: Markerless Hand-Tracking for Computer Aided Design. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 549–558. https://doi.org/10.1145/2047196.2047269

[58] Andrew M. Webb, Andruid Kerne, Zach Brown, Jun-Hyun Kim, and Elizabeth Kellogg. 2016. LayerFish: Bimanual Layering with a Fisheye In-Place. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces* (Niagara Falls, Ontario, Canada) *(ISS '16)*. Association for Computing Machinery, New York, NY, USA, 189–198. https://doi.org/10.1145/2992154.2992171

[59] Daniel Wigdor, Hrvoje Benko, John Pella, Jarrod Lombardo, and Sarah Williams. 2011. Rock & Rails: Extending Multi-touch Interactions with Shape Gestures to Enable Precise Spatial Manipulations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 1581–1590. https://doi.org/10.1145/1978942.1979173

[60] Christian Winkler, Ken Pfeuffer, and Enrico Rukzio. 2012. Investigating Mid-air Pointing Interaction for Projector Phones. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. ACM, New York, NY, USA, 85–94. https://doi.org/10.1145/2396636.2396650

[61] Anusha Withana, Makoto Kondo, Yasutoshi Makino, Gota Kakehi, Maki Sugimoto, and Masahiko Inami. 2011. ImpAct: Immersive Haptic Stylus to Enable Direct Touch and Manipulation for Surface Computing. *Comput. Entertain.* 8, 2, Article 9 (Dec. 2011), 16 pages. https://doi.org/10.1145/1899687.1899691

[62] Siju Wu, Amine Chellali, Samir Otmane, and Guillaume Moreau. 2015. TouchSketch: A Touch-based Interface for 3D Object Manipulation and Editing. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (Beijing, China) *(VRST '15)*. ACM, New York, NY, USA, 59–68. https://doi.org/10.1145/2821592.2821606

[63] Haijun Xia, Ken Hinckley, Michel Pahud, Xiao Tu, and Bill Buxton. 2017. WritLarge: Ink Unleashed by Unified Scope, Action, & Zoom. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 3227–3240. https://doi.org/10.1145/3025453.3025664

[64] Min Xin, Ehud Sharlin, and Mario Costa Sousa. 2008. Napkin Sketch: Handheld Mixed Reality 3D Sketching. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology* (Bordeaux, France) *(VRST '08)*. Association for Computing Machinery, New York, NY, USA, 223–226. https://doi.org/10.1145/1450579.1450627

[65] Baoxuan Xu, William Chang, Alla Sheffer, Adrien Bousseau, James McCrae, and Karan Singh. 2014. True2Form: 3D Curve Networks from 2D Sketches via Selective Regularization. *ACM Trans. Graph.* 33, 4, Article 131 (July 2014), 13 pages. https://doi.org/10.1145/2601097.2601128

[66] Ka-Ping Yee. 2004. Two-Handed Interaction on a Tablet Display. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1493–1496. https://doi.org/10.1145/985921.986098

[67] Dongwook Yoon, Nicholas Chen, and François Guimbretière. 2013. TextTearing: Opening White Space for Digital Ink Annotation. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. ACM, New York, NY, USA, 107–112. https://doi.org/10.1145/2501988.2502036

[68] Dongwook Yoon, Nicholas Chen, François Guimbretière, and Abigail Sellen. 2014. RichReview: Blending Ink, Speech, and Gesture to Support Collaborative Document Review. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 481–490. https://doi.org/10.1145/2642918.2647390

[69] Robert Zeleznik, Andrew Bragdon, Ferdi Adeputra, and Hsu-Sheng Ko. 2010. Hands-on Math: A Page-based Multi-touch and Pen Desktop for Technical Work and Problem Solving. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. ACM, New York, NY, USA, 17–26. https://doi.org/10.1145/1866029.1866035