



**HAL**  
open science

# Exploiting Physical Referent Features as Input for Multidimensional Data Selection in Augmented Reality

Gary Perelman, Marcos Serrano, Emmanuel Dubois

► **To cite this version:**

Gary Perelman, Marcos Serrano, Emmanuel Dubois. Exploiting Physical Referent Features as Input for Multidimensional Data Selection in Augmented Reality. ACM Transactions on Computer-Human Interaction, In press. hal-04323637

**HAL Id: hal-04323637**

**<https://hal.science/hal-04323637v1>**

Submitted on 18 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Exploiting Physical Referent Features as Input for Multidimensional Data Selection in Augmented Reality

Gary Perelman

IRIT, University of Toulouse III, Toulouse, France. gary.perelman@irit.fr

Marcos Serrano

IRIT, University of Toulouse III, Toulouse, France. marcos.serrano@irit.fr

Emmanuel Dubois

IRIT, University of Toulouse III, Toulouse, France. emmanuel.dubois@irit.fr

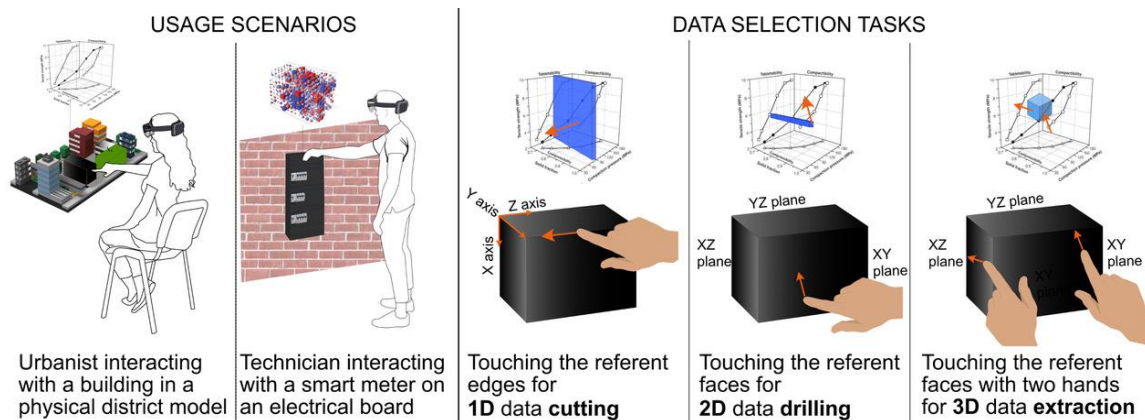


Figure 1. Our interaction setting: a user interacts with a physical referent, either a physical model of a city district or an electric board. A space-time cube, displayed above the referent using a Head-Mounted Display, shows the energy consumption data. The user can perform 1D, 2D or 3D data selection tasks by touching the referent edges or faces.

Embedding data into the physical environment using augmented reality (AR) is a practical approach for data visualization as it offers a large and flexible display space on or around the physical referent, i.e. the physical object to which the data is related. Yet current interaction in such context is often performed using cumbersome dedicated devices, tiring mid-air gestures or awkward on-body input. In this paper, we investigate the use of the physical referent itself as a support for input interaction with an embedded space-time cube (STC) representation. Hence, we first identify the most promising mappings between the physical features of the referent (edges, faces, corners) and the STC dimensions. Then, we design three data selection techniques using the physical referent and compare them to mid-air gestures when performing selection tasks on the STC. Overall, our work demonstrates that using the physical referent to support input interaction with embedded data representations is an efficient and comfortable approach for data selection in standing and sitting situations.

CCS CONCEPTS • **Human-centered computing** → **HCI design and evaluation methods**; • **Human-centered computing** → **Mixed / augmented reality**;

**Additional Keywords and Phrases:** augmented reality, physical referent, interaction techniques

## 1 INTRODUCTION

Smart environments are now common at different scales, from smart cities [1][25][33] and smart industries [100] to smart homes [9], thanks to the spread of sensing devices collecting large volumes of data, such as energy consumption data [72][84]. The in-situ access to such data is crucial to help stakeholders, workers or citizens in decision-making processes [33][72] on-the-go or within their physical environment [51]. This is particularly relevant for field technicians (e.g. to address energy consumption problems) or when people collaborate around a physical object (e.g. urban planners around a physical district model) to better understand the observed data in its physical environment.

A practical solution to conduct such data access is to display the data within the physical environment using augmented reality (AR), an approach also known as embedded data visualization [51][118]. This approach is beneficial for visualizing a large amount of data, as it reduces information clutter thanks to the almost unlimited AR display space [47]. It also facilitates data retrieval and understanding [78] by offering a physical anchor to data visualization. Indeed, such data can be visualized on or around **the physical referent**, i.e. the physical object which aggregates the data (e.g. the electricity meter or other sensing devices) or to which the data is related (e.g. a building on a physical model). And yet, despite the benefits mentioned above, the interaction with this embedded data is challenged by the lack of well-suited input interaction modalities.

Previous work has explored three leading solutions to interact with digital data displayed in AR: using dedicated devices, mid-air interaction or on-body interaction. First, using dedicated objects such as physical controllers [27], and surrounding devices such as smartphones [88] or tangibles [121] provides a rich interaction vocabulary and facilitates interaction thanks to tactile cues. However, using such auxiliary objects requires the user to always have them at hand, which can be impractical to interact with embedded visualizations on the go. Second, head-mounted displays (HMDs) natively detect mid-air interaction, thus limiting the need to track external devices. They also allow for direct 3D interaction with digital data representations, providing the user with a natural and fluid interaction. However, mid-air interaction can be tiring [48][50], and users must know mid-air gestures beforehand. Third, on-body interaction [89] offers a rich interaction vocabulary available everywhere. While it is efficient for navigation tasks [96] or menu interaction [4], we did not find previous work focusing on data selection in an AR visualization. Furthermore, on-body interaction can raise social acceptability issues [53]. To sum up, we are still missing an efficient interaction approach to perform the data selection in embedded visualizations.

In this paper, we explore how to provide the user with the appropriate interaction techniques to select data in embedded visualization. Since previous works have demonstrated the benefits of using a physical referent for embedded data visualization [51][118], we investigate how to take advantage of the same physical referent to support the input interaction with the embedded data. Indeed, using the physical referent as an input device provides a physical gesture delimiter, thus 1) removing the need for interaction modes or additional input devices, 2) overcoming the limitations of mid-air interaction (fatigue, inaccuracy), and 3) facilitating eyes-free input, thanks to the use of its physical features, to keep the visual attention on the data. Besides, embedding input interaction on the physical referent will contribute to a more fluid and contextualized visualization and manipulation of its data.

While this approach may be suitable for a wide variety of data visualizations, as a first step in this paper, we focus on selecting multidimensional data in a space-time cube (STC) visualization displayed above the physical referent in AR. Indeed, STC visualization is commonly used in data analysis [6][37]. We considered three fundamental data selection tasks [7] (1D orthogonal cutting, 2D orthogonal drilling and 3D point extraction) needed to perform any more complex tasks (such as filtering or comparing multiple datasets). These selection tasks are also valuable for various visualizations (e.g. 1D timelines, 2D graphs).

We first systematically explore how to use the physical features of a physical referent (edges, faces, corners) to perform selection tasks on the STC. To do so, we explore in a preliminary study the user preferences for interacting with the physical referent and identify the most promising mappings between the physical features of the referent and the STC data dimensions. Then, we design a set of interaction techniques using the physical referent to select data in the STC: one technique is entirely based on touch interaction on the physical referent, and two techniques combine touch interaction on the physical referent with mid-air gestures. We compare these techniques with full mid-air interaction. Then, we iterate on the design of our techniques to enhance their performance for the 3D extraction task and compare them to full mid-air interaction. Finally, we evaluate if the previous results, measured for a horizontally fixed referent with a **sitting** user, are still valid when users are **standing** in front of a horizontally fixed (e.g. on a table) or a vertically fixed (e.g. on a wall) physical referent.

Our contributions are as follows: 1) the exploration of the mapping between the physical referent features, STC dimensions, input modes and data selection tasks (1D cutting, 2D drilling and 3D extraction) and 2) the design and evaluation of a set of interaction techniques for these data analysis tasks when the user is seated or standing in front of a vertically or horizontally fixed physical referent.

## 2 RELATED WORK

Our work is related to embedding data into the environment. It focuses on combining the advantages of using augmented reality to display data and using a physical referent to support input interaction in this context. In this section, we review previous works related to these topics.

### 2.1 Embedded Visualization

Regarding the significant topic of embedded visualization, we first focus on the embedded visualization paradigm and how it can be implemented in our context. Then, we discuss the different data representations used in embedded visualizations.

Embedding data into the physical environment allows users to engage through senses beyond vision [114], thus allowing eyes-free input interaction with data [52][65], enhancing visualization tasks [28] and facilitating memorization and proprioception [107]. Willet et al. [118] proposed a design space to characterize different levels of embedded data representation along four categories: non-situated, situated, embedded physicalization and embedded visualization. Non-situated and situated visualizations facilitate the data visualization by extending the screen real-estate through large devices (wall display, tabletop) [87] or multiple devices [11] [41][45]. However, these approaches still limit the input interaction to mouse or touch input and the visualization to 2D screens. Embedded physicalization relies on a physical artefact that can be directly actuated or modified (e.g. light) [42][94][106][108]. However, these solutions mainly focus on data visualization rather than input interaction and are currently limited by the technology. On the one hand, static physicalizations [68][69][122] hinder the visualization of multiple datasets as well as the analysis amongst numerous dimensions, the use of filters, etc. On

the other hand, dynamic physicalizations [92][94][108] need complex actuated systems, which can be expensive and difficult to generalize to multiple datasets and use cases. Embedded visualizations, the last category introduced by Willet et al. [55][116], depict in-between situations where the data is projected on or near a physical artifact [25][77]. This approach is particularly well suited for in-situ data visualizations.

Previous work showed that augmented reality with a see-through Head-Mounted Display (e.g. the HoloLens) enhances embedded data visualization thanks to the benefits of immersive cues [115]. Layering data visualizations directly in the user's environment facilitates data understanding by allowing spatial exploration [10][32][36][47]. In addition, mapping data on or around physical objects enhances data re-contextualization and understanding [78]. For example, in DataHop [47], data is spread over multiple physical locations in the user's environment, each spot corresponding to a specific data analysis. To enhance the use of the HMD, some solutions combine it with another device (e.g. a smartphone or a smartwatch) [11][41][45][88]. In this context, the external device can serve as an anchor [41][45] for data visualization, as an input device [88] or as a complementary screen [123] to visualize data differently (e.g. 2D chart).

Now, to create the rendering of embedded data when facing an extensive collection of data, multiple forms of embedded visualizations have been explored [21][31][56][67]. Among them, the Space-Time Cube (STC) [6][56] is particularly relevant. It can be displayed above a referent (e.g. map, model) to facilitate the visualization of geo-temporal data frequently used in many domains such as smart-cities [38] or dynamic networks [8]. In these cases, the spatial relationship between data points and their physical origin is beneficial (e.g. for identifying groups), and their spatial exploration in augmented reality facilitates the data understanding [78].

Overall, using AR to display a STC near the physical referent is a well-suited solution for embedded visualization, allowing the user to be aware of his environment while visualizing the data representation. However, in this context, we still lack efficient interaction techniques to manipulate the data representation, as underlined by Ens et al. [35]. They identified input interaction as one of the key challenges of immersive analytics and, more specifically, for embedded data visualizations.

## **2.2 Input interaction for embedded data visualizations**

In this section, we synthesize three input interaction paradigms for embedded data visualizations: mid-air interaction, interaction with dedicated devices (e.g. remote controllers, smartphones) and on-body interaction.

The use of mid-air gestures [22][47][49] remains the norm for interacting with embedded data visualizations in AR. Mid-air interaction is almost always available and directly detected by the HMD, thus limiting the need for external tracking devices. But mid-air interaction also induces physical fatigue, a limited gesture vocabulary and low discoverability. Previous research proposed to reduce the fatigue and imprecision of distant pointing by using virtual proxies [17][83], i.e., replicating a portion of a remote space near the user. These solutions provide a nearby area, either a plane [17] or a sphere [83], where the user can perform the mid-air gestures. These gestures are then replicated in a distant environment, thus enhancing precision and limiting the need to maintain the arm in mid-air for an extended period. However, these solutions do not answer the discoverability issues and remain limited by the gesture detection capabilities of the HMD, leading to a restricted interaction vocabulary.

The second primary interaction approach consists in using physical wireless controllers such as the HTC Vive controllers [27][82][105] or a combination of devices [3][20][70][117]. In this context, the external device can be used for input [27][70][82][105]: with ImAxes [27], a remote controller is used to manipulate data axes and combine them to generate charts. Similarly, in TrackCap [70], a smartphone is used as a 6 degrees of freedom

controller providing touch input. Alternatively, the external device can also be used as output [3][20][117]. In SymbiosisSketch [3], a tablet is used to visualize drawings and properties of 3D objects displayed in augmented reality. While these solutions provide a physical support to perform the interaction and extend input and output capabilities, the external controller or device still needs to be available near the user, limiting the possibility of interacting on the go with embedded data around a physical referent.

Finally, the third interaction approach consists in using the body as input, i.e. on-body interaction [4][63][89][96]. This approach takes advantage of the permanent availability of the user's body to extend the input vocabulary. Among the body parts explored in previous work, Azai et al. [4] used the forearm as a touch surface to configure a 3D object displayed in augmented reality. Saidi et al. [89] exploited it to facilitate data manipulation in immersive visualizations. While this body part has been the most explored due to its availability and large surface [63][89], other body parts, such as the face [96] or the wrist [97], have been shown to facilitate interaction with HMDs. However, efficiently tracking on-body gestures is still an open issue which often requires using external devices since most body parts are not within the field of view of the HMD camera [4]. Furthermore, on-body interaction often raises social acceptability issues [53].

Given the limitations of existing input interaction approaches, we suggest focusing on the opportunistic use of the physical referent, already available in the user's environment and which can be tracked using a see-through HMD. This approach is similar to tangible user interfaces (TUI) as it exploits the physical properties of the referent. We therefore synthesize existing work on TUI in the following section.

### **2.3 Tangible User Interfaces**

Tangible User Interfaces (TUIs) use a physical object as input [34][86][121] to interact with data, virtual objects or applications. Such an approach is often used to provide more degrees of freedom [43] or to allow eyes-free interaction [52] and thus facilitates navigation and selection in 3D visualizations [13]. Tangible interaction also takes advantage of proprioception, enhances memorization [107] and is easier to apprehend [98]. Previous works [12][109] showed that tangible interaction is faster than tactile or mouse interaction in specific contexts. For example, Besançon et al. [13] compared tactile, tangible and mouse interaction to perform a 3D docking task. They found that tangible interaction was faster and more enjoyable than tactile or mouse interaction.

Designing effective tangible interaction techniques needs to carefully consider the shape and size of the manipulated tangible [81][121]. Previous works showed that the correlation between the tangible object and its digital counterpart (e.g. a cursor, a menu, a virtual object) is highly relevant [58]. As a result, some works focused on providing the tangible with specific shapes such as spheres [34], cubes [2][38] or even reconfigurable shapes to enhance interaction. For example, Zhao et al. [121] developed a tangible object based on a robotic assembly whose shape could adapt to accurately match the form of the virtual object to manipulate and offer a more realistic grasping. Another property to consider is the affordance of the tangible, i.e. how its form factor suggests the gripping, manipulation and purpose [71].

Tangibles have already been used to interact with data visualizations. In such context, they can be used to physicalize the axes of the data representation as explored in Embodied axes [26]: the physical object representing an axis possesses actuators and buttons to configure the axis scale factor, minimum and maximum. Smiley et al. [101] used a similar approach with several tangibles representing the different axes: this allows to generate various data visualizations depending on the spatial relationship between the tangibles. Alternatively, tangibles can offer a physical delimiter such as in CAPTIVE [22], where a wireframe cube delimits the data representation and

interaction space. Manipulating the cube (e.g. rotate) allows to interact with the data visualization while interacting inside the wireframe enables manipulating the data itself.

To conclude this synthesis on tangibles, previous approaches fit into two categories: 1) tangibles based on simple shapes like cubes or spheres [2][34][38] which are usually bound to simple tasks, thus requiring multiple tangibles for more complex scenarios or 2) complex tangibles [26][101] using actuators, buttons and custom shapes which can be used for a wide range of tasks but are less commonly found in the user's environment (as they are specifically designed for a set of tasks), thus limiting their generalization.

Furthermore, most of these tangibles are intended to be grabbed by the user [22][38][34][121]. But this is not always possible in the case of embedded data visualizations because physical referents can be attached, stuck, wired, etc. In this work, our approach differs from the usual tangible approach: we propose to explore how to take advantage of a physical referent that cannot be grasped and freely manipulated as support for input interaction with the data to which it is related.

### **3 USAGE SCENARIO AND DESIGN CONSIDERATIONS**

To better contextualize our work focusing on using a physical referent as an input support for interacting with embedded data, we first present two usage scenarios of multidimensional data selection in AR with a physical referent as input. The first scenario focuses on a field technician conducting on-site work, while the second presents a seated urban planner working around a physical district model. We then derive a set of design considerations for our interaction approach, including the data representation, the user tasks and the interaction techniques.

#### **3.1 Usage scenarios**

In the following scenarios, we used the gender-neutral pronoun "they" to refer to our user.

##### *3.1.1 Field technician performing on-site work*

Chris is an electric technician at a smart campus [76][110]. Most of their work takes place on-site, and interacting with data directly in-situ allows them to identify the problems and solve them all at once, as illustrated by the following example. As Chris operates a maintenance task in a building, they realize that the room temperature is too high. To evaluate whether this anomaly is a one-time occurrence or has been going on for a few days, they grab a Head-Mounted Display (HMD) and touch a smart controller (Figure 1) used to sense and automatically adjust the parameters of the room. This triggers a data visualization of the room's temperature per day and hour. By touching the sensor's corners, faces or edges, they discover that this problem has happened every morning for a few days. Thus, they control each radiator with their technician tools. The issue is that one radiator thermal sensor is out of order. They deactivate the defective radiator, which temporarily solves the problem.

*Similar related contexts:* The situation depicted in this scenario can be extended to different contexts, such as at home for monitoring IOT sensors[99]. At home, the user could explore its energy consumption by interacting on its electric meter or on the panels of the radiators, as illustrated by the different existing smart sensors in the figure below. This scenario could also be applied to an industrial context (Industrial Internet of Things [100]) where one technician goes through the production chain of a factory to detect anomalies. In this context, the command control panel of each machine could be used as physical input.

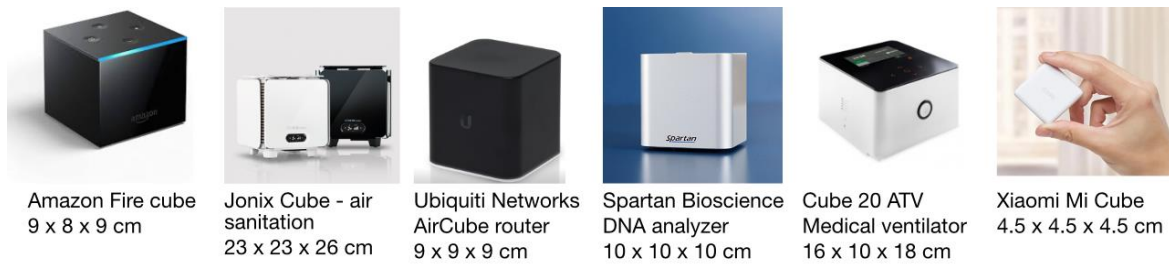


Figure 2. Examples of small-sized cubic-shaped physical devices populating our environment and collecting or generating heterogeneous data, along with their dimensions.

### 3.1.2 Urban planning and mobility

The following scenario is based on *recent* projects focusing on urban planning for smart cities [25][110]. Several stakeholders, such as elected officials, energy technicians, urban promoters and architects, collaborate to design an appropriate urban planning solution. To facilitate communication and interaction among them, the project explores using a shared physical model linked to a data visualization generated through a simulation platform running in the background. This physical model represents the city district's various buildings, roads and urban furniture. At the same time, the simulation generates data such as parking spot occupation or public transport load. Dominique, an urban planner, wants to explore the impact of adding a building on urban mobility. They grab a Head-Mounted Display (HMD), sit on a chair near the table on which the physical model is placed and touch the four floors car park (Figure 1). This triggers the display of a data visualization above the physical model of the car park (average number of free parking spots per hour and day over a week). They then position a new building in a vacant lot: touching this new building gives them access to a possible range of values regarding the number of flats or residents. While modifying them, Dominique can study the impact on the available parking spots by day and hour in the representation displayed above the car park. They can then derive recommendations for street parking or public transportation development, according to the size of the new building.

*Similar related contexts:* this scenario can be extended to the context of crisis management [30]: stakeholders share a physical model of a city district to favor their collaboration. This model easily identifies buildings threatened by a fire over time and estimates the number of people at risk. Such a physical overview would help predict the needed resources and support the decision-making process. While a simulation running in the background generates the data in our scenario, using real data collected from smart sensors or smartphone apps would also be possible. This scenario could also be of interest in an industrial context. A physical model of the factory could be used in a control room to reorganize the production hall (e.g. move the machine's position and explore the resulting theoretical production efficiency) or monitor the real-time parameters of the production chain.

### 3.1.3 Summary of our usage scenarios

These two scenarios illustrate two different situations in which accessing data on a physical referent is of interest. In both scenarios, the physical context of the data can easily be perceived and considered for the data exploration. Furthermore, in the first scenario, the user could discover and explore data while at the data source location. This offers the advantage of allowing the user to analyze and solve any problem simultaneously. In the second scenario, the interest of accessing data on a physical model is to understand better the spatial relationship between the buildings or other physical resources and their associated data: for instance, the impact of a building's number of



floors on the available parking spots nearby. Both scenarios depict a user interacting with a cubic-shaped physical referent (a sensor or a building model). Still, they differ regarding the context: a user standing and interacting with a physical referent attached to a vertical surface in the first scenario or a user seating and interacting with a referent placed on a horizontal table in the second scenario. In the following section, we use these two scenarios to inform and motivate a set of design considerations for our approach.

## 3.2 Design considerations

### 3.2.1 Data representation

Datasets may contain numerous dimensions [72][84] over space and time (energy consumption, temperature, population density, etc.). Hence, in our context, we display the data using a space-time cube (STC) representation [6][56]. In our scenarios, such a representation is particularly relevant to support comparisons among days of the week or periods of the day. A STC would also be particularly well suited to represent geographically situated information, such as comparing consumption among different buildings located in various places on a map.

### 3.2.2 Interaction with a data visualization

To interact with the STC data, we refer to a set of 36 well-established analytical tasks [6][12][29]. We identified three preponderant sets of tasks from these analytical tasks: data selection, chart manipulation (rotating/scaling the STC) and data filtering. In this paper, we focus on the selection only, which is required before performing any other analytical task. We identified three different selection tasks (see Figure 3): orthogonal cutting (hereafter named 1D cutting), orthogonal drilling (2D drilling) and point extraction (3D extraction). 1D cutting and 2D drilling tasks can be applied respectively to the STC's three axes (X, Y and Z) and three planes (XY, XZ, YZ). Hence, we considered seven tasks: 1D cutting on X, Y and Z axes, 2D drilling on XY, XZ and YZ planes and 3D extraction for a unique point with XYZ coordinates. While these tasks are designed for a three-dimensional STC, 1D cutting and 2D drilling selections can be used on data visualizations such as 2D charts or 1D timelines. They also correspond to generic tasks in which only one or two dimensions need to be controlled (selecting a parameter among several, adjusting a slider position, selecting one colour in a colour map, etc.).

We illustrate the three selection tasks using our scenarios and their cubic-shaped physical referents. The STC displays the data (energy consumption or available parking lots, Y axis) according to an hourly scheme (X axis) for each day of the week (Z axis). A 1D cutting selection would correspond to the selection of the data value on Monday (Z axis cutting) or over the week between 10 and 11 am (X axis cutting). A 2D drilling would allow the user to visualize the data value on Tuesday between 8 and 9 am (XZ plane drilling). Finally, a 3D extraction would allow focusing on a specific value, e.g. a peak value on Wednesday between 4 and 5 pm.

### 3.2.3 Input modes

To explore a large variety of solutions, we consider three different input modes on or around the physical referent (see Figure 3): 1) the regular mid-air interaction used in AR (e.g. HoloLens), 2) direct touch interaction on the physical referent [19][120], or 3) indirect tangible interaction, by holding an additional tangible brick, external to the referent, either in mid-air or in contact with the physical referent (i.e. similar to Touché [91]). We did not consider holding the physical referent; in our scenarios, the referent is fixed to a wall (electric meter) or a table (building of the physical model).

Our rationale for focusing on these input modes is the following: Mid-air input (hand position or gestures) is the most common and known approach to interact in AR, as well as the default input on commercial devices such as HoloLens, making it a suitable baseline. Touch interaction on the physical referent is the most prominent and direct solution: one or several fingers can touch the edges and faces of the physical referent. Finally, the tangible brick offers more degrees of freedom than tactile interaction on the physical referent. We consider it as a probe, assuming that in some specific cases, users could opportunistically grasp a nearby everyday life object (e.g. a USB key, a card, a lighter or a fluo marker). Furthermore, tangible interaction is easy to apprehend and can potentially greatly enrich the user’s input vocabulary with the object’s orientation and position. Previous works exploring using a cubic shape as a unique but universal tangible widget or cursor [2] inspired our design.

Finally, we considered using these input modes with one or two hands, except for the mid-air interaction mode, which was performed with one hand only. Indeed, a single hand moving in mid-air corresponds to a single point in 3D, which can be used to select one point inside the 3D STC.

### 3.2.4 Physical referent and its physical features

This paper focuses on a cubic-shaped physical referent because it is a widespread and generic shape. As illustrated in our scenarios above, it also adapts well to a variety of scenarios and can represent different real-world objects such as a 3D printed building, an electricity meter, an internet box, various smart devices (see Figure 2) or even a display case protecting a model (museum, airports). The Discussion section further elaborates on how our approach could generalize to other shapes.

Such a cubic shape fixed on a surface offers a rich space of 17 physical features for interaction (four corners, eight edges and five faces) after removing the face in contact with the surface and its inaccessible edges and corners (see Figure 3). Not all these physical features are relevant for interaction: for instance, when the physical referent is fixed on a horizontal (respectively vertical) surface, the face and edges that are behind (respectively below) the physical referent are probably less reachable, and the features on the left can be difficult to touch with the right hand.

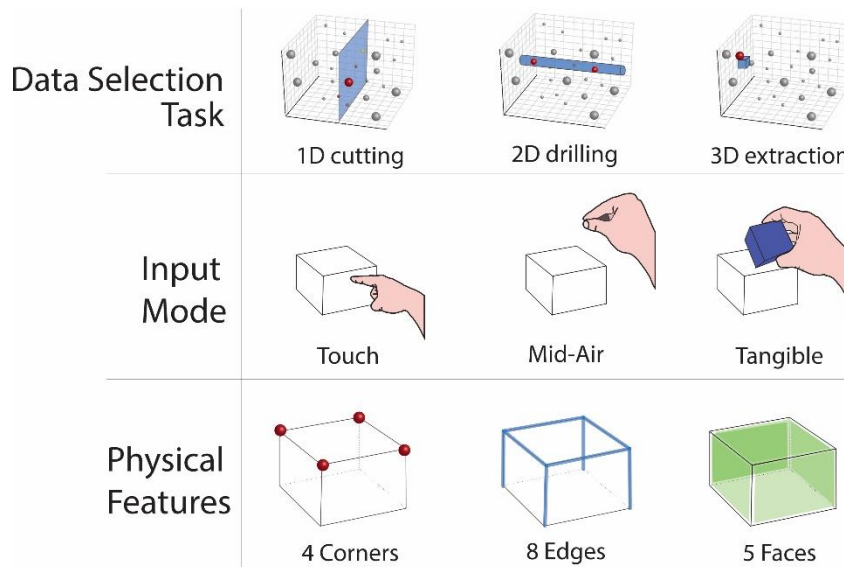


Figure 3. The dimensions of our exploration: Data selection tasks, Input modes and Physical features of the referent.

Based on these design considerations, we explore in the following studies how users perform the three identified tasks (1D cutting, 2D drilling, 3D extraction) through the possible combinations of input modes (mid-air, touch, tangible), physical features (surfaces, edge, corner) and fingers/hands.

#### **4 PRELIMINARY STUDY: MAPPING PHYSICAL FEATURES WITH TASKS**

This preliminary study aimed to identify which physical features (edges, corners, faces) of the physical referent are the most appropriate to use when interacting with a digital space-time cube (STC) and how each input mode can contribute to this goal. To this end, participants were asked to propose, with each input mode, a way to perform data selection tasks on the STC using the physical referent.

##### **4.1 Study Description**

We recruited 12 participants, all with a background in computing science (28.6 years old on average, 5 females, 7 males). Participants did not receive any compensation for the study. Participants were seated in front of a desk, on which was placed a physical model representing a cubic-shaped object fixed (using tape) on a horizontal surface (e.g. a physical model of a city district). The dimensions of the physical model were 10cm x 10cm x 10cm, similar to existing cubic-shaped smart devices (see Figure 2). The associated data was displayed in a STC above the physical referent using a HoloLens.

We asked participants to explore the physical features of the referent (faces, edges, corners) and imagine how they would perform each of the seven data analysis tasks (1D Cutting in the X, Y and Z axis, 2D drilling in the XY, YZ, XZ planes, 3D extraction: see section 3.2.2) the most naturally and comfortably with one or two hands. Each participant repeated this exploration three times: each time, the experimenter instructed them which of the three input modes (mid-air gestures, direct touch and indirect use of an additional tangible brick) had to be used, potentially in combination with another one. It ensured that the participants would consider our three input modes. They had to express the chosen solution verbally, and the input modes' order was counterbalanced across the participants.

In the case of the tangible brick, and to account for the possible effect of its size [58], we 3D printed three different dimensions (2, 3 and 6 cm) and let users change the brick at any time during the experiment. Users prefer tangibles they can hold with one hand [81], so all of our tangible bricks were small enough to fit in the users' hand. Our tangible could be used like Conté [112] by touching the physical referent with it, in which case a smaller tangible brick might be attractive. But it could also be directly touched by the user, in which case a larger tangible is more appropriate as it may facilitate precise interaction.

Finally, we asked participants not to use the same action for two different tasks. Hence, their final set of mappings represented a viable global approach. As a result, we collected 3 (input modes) x 7 (selection tasks) x 12 (participants) = 252 mappings. There was no time limit for the exploration and the total duration of the experiment took approximately 45 minutes.

At the end of the experiment, users had to rank their preferred input modes overall. We also recorded which physical features (tangible brick or referent) participants used and their mapping to the STC dimensions. As our goal was to let users freely explore the physical features, we did not implement the input modes or any system feedback at this stage: users had to imagine and physically execute the actions or gestures that would lead to a data selection on the STC, without observing the effect on the STC.

## 4.2 Results

To analyze the results of this study, we did not perform any statistical treatment. Instead, we counted the number of participants who adopted the same or similar solutions to identify the actions (associated to each of the seven task) proposed by most participants.

### 4.2.1 Input-mode usage

**Input-mode preferred:** The results on ranking input modes established that the preferred interactions involved touching the physical referent directly (11/12) rather than using the intermediate tangible brick or mid-air interaction. Users commented positively on the direct touch solution. They found touch interaction intuitive (P4) and convenient (P7). P11 confirmed by saying: *“Touch interaction is rather intuitive; I feel like I’m touching the faces of the space-time cube”*. On the contrary, they reported that *“Using the tangible brick for a task like pointing is too complex (P1)”*, meaning that pointing is too simple for justifying the rich interaction vocabulary provided by the tangible brick. The only participant who preferred to use the external tangible brick stated: *“I have more possibilities with the tangible brick than with simple touch (P9)”*. Interestingly, none of the participants mentioned the mid-air interaction as their preferred solution.

**Proposed solutions with direct touch:** Users always performed 1D cutting and 2D drilling with a single hand (12/12). Considering 3D extraction, 4/12 participants proposed single-handed solutions (using multiple fingers or different physical features) and 8/12 proposed bimanual interaction (3 using both hands on the physical referent and 5 using the dominant hand on the physical referent and the non-dominant hand in mid-air).

**Proposed solutions with the tangible brick:** The first interesting result is that most of the participants (8/12) used the 3cm version. Three preferred the 2cm version and only one the 6cm version. Then, users mainly chose to perform translations with the tangible brick in mid-air (9/12) to move the virtual cursor. One participant performed rotations with the tangible brick in mid-air, and two chose to use the tangible brick as a “pen” directly on the physical referent. Participants discarded exploiting a more developed vocabulary based on the tangible brick, as confirmed by their comments: *“The more physical features I combine, the more I feel lost. (P1)”*; *“Using the tangible brick on the model is too constrained; I prefer using it mid-air (P8)”*.

**Proposed solutions with mid-air interaction:** All participants proposed similar solutions where the dominant hand moves in mid-air to manipulate a virtual cursor inside the STC. Interestingly no participant proposed using a ray-casting metaphor.

### 4.2.2 Preferred physical features

We now report which physical features of the referent were most often used for each task, as illustrated in Figure 4.

For 1D cutting, participants only used the edges of the model (12/12) (top edge of the front face for X (10/12), left edge of the front face for Y (7/12) and left edge of the top face for Z (7/12)). For each axis, they only considered one other solution: for the X axis, two participants used the bottom edge of the front face (i.e. they touched the bottom of the front face as the bottom edge is not reachable); for the Y axis, five participants used the right edge of the front face; for the Z axis five participants used the right edge of the top face. In addition, results also indicate that when users chose to use the left edge of the front face for Y, they also used the left edge of the top face for the Z axis (and respectively, the right edge of the front face alongside the right edge of the top face). These results confirm that users use 1-dimensional physical features to perform 1-dimensional selection tasks on the STC.

For 2D drilling, participants mainly used the front face (XY) (11/12), the right face (YZ) (9/12) and the top face (ZX) (11/12) of the physical referent. These results are explained by the fact that accessing the right face is easier with the dominant hand than the left face, as stated by P1, P4 and P7. The remaining participants used the left face (XY) (1/12), the left (2/12) and front (1/12) face for YZ and the right face (1/12) for ZX.

For 3D extraction, participants (7/12) mainly used the front face (XY) for XY movements and the top and left faces for Z movements. Other participants (5/12) used the front face for XY and their non-dominant hand in mid-air for the movements along the Z axis. As we requested users to propose distinct solutions for each task; using the top face for Z movements was disambiguated from using it for the 2D drilling task by using two fingers simultaneously (4/7) or through the use of bimanual interaction (3/7).

The physical referent corners were only used for task selection (e.g. switching from 2D drilling to 3D extraction) (2/12). Other participants didn't think of a switching mechanism between 2D drilling and 3D extraction as they provided distinct interaction techniques for the two tasks.

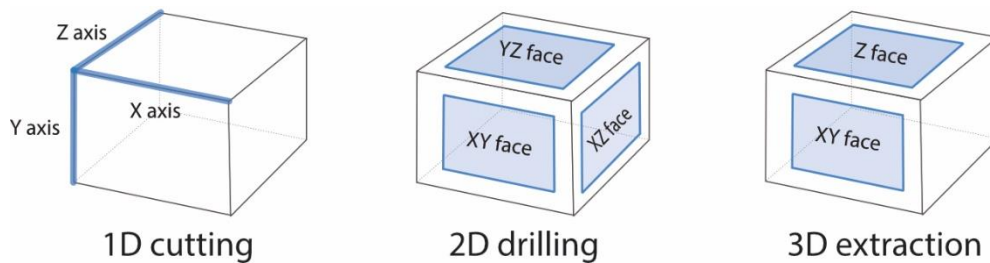


Figure 4. Preferred physical features for each data analysis task according to our study results.

### 4.3 Summary

To sum up, our study reveals that users prefer direct touch interaction on the physical referent (mono or bimanually) and that mid-air interaction, despite being the most common and known interaction, was never cited as the preferred interaction mode. Regarding mono or bimanual interaction on the physical referent, participants preferred mainly bimanual interaction (with the non-dominant hand in mid-air) and mono-manual touch on the physical referent.

Regarding the physical features, participants prefer using the physical edges of the referent for 1D cutting and its faces for 2D drilling and 3D extraction (with a mode-switching mechanism to disambiguate the use of the face when required).

Results also reveal that using an external tangible brick is too complex for the selection tasks proposed in this preliminary study. Consequently, in accordance with our initial motivation to adopt a device-less approach (see sections 1 and 2.3), we discarded this input mode in the following studies. We believe that adding an external tangible brick with too many degrees of freedom was too cumbersome: there were 7 tasks to perform while the tangible brick offered more than 20 possible interactions.

Based on these results, we design and implement four interaction techniques in the following section based on touch only, mid-air only (the baseline), or a combination of both in mono- and bi-manual form.

## 5 INTERACTION TECHNIQUES FOR EMBEDDED DATA SELECTION

In this section, we present the design of a set of interaction techniques for 1D drilling, 2D cutting and 3D extraction on a STC based on the preliminary study results. After introducing the overall design rationale of the techniques, we first present a mid-air only technique (the baseline, not using the physical referent) and then our three referent-based techniques in detail. We compare the performance of these four techniques in the following study (section 6).

### 5.1 Design rationale

We structured the design of our interaction techniques around four design considerations.

*Task change:* One of the advantages of having multiple physical features on the referent is to allow for an implicit task change. For instance, the user should be able to switch from 1D cutting to 2D drilling seemingly. To make this implicit task change possible, we can't use the same input gesture (i.e. user action on a physical feature) for two different tasks.

*Input range:* As illustrated in our usage scenarios, each axis should allow selecting a wide range of values (e.g. timelines, sensor values, etc.). Thus, we used a 1-100 range on each axis. Informal tests performed with 4 participants revealed that selecting with such a wide range was imprecise when using mid-air or touch interaction. Instead of limiting the range of data to 10 values and hence limiting the external validity of our techniques, we decided to rely on a dual precision approach as commonly used in the literature [24][61][104][111]. We decomposed the data selection into two steps: a first step in which users select the **tens** digit and a second in which users select the **unit** digit. Thus, both steps require the same dexterity as users must select one value among ten.

*Selection and validation:* Given the dual precision approach, we needed to design a validation mechanism for the tens and units steps. We used the same approach for all the techniques, to ensure consistency during our experimental comparisons: a dwell time of 500ms [74][124], activated when the user's cursor stands still. After this time, the tens range selection is validated and the units step starts. The user can cancel the selection and start a new one during both steps. However, once the digit step starts, there is no way to return to the tens step.

*Input modes:* Given the considerations mentioned above and the results of our preliminary study, we designed three interaction techniques involving some form of direct touch interaction on the physical referent. We name our techniques as a product of the input mode (IM) used for each selection step, as in previous work [75] :  $\langle IM_{tens} \times IM_{units} \rangle$ . This naming convention explicitly mentions which technique was used for the tens and the units steps of the selection process, as these techniques may differ in some cases. If both steps use the same technique, we used the square sign (e.g.  $Touch^2$ , meaning  $Touch \times Touch$ ). By default, the input mode is performed with the dominant hand. If the interaction requires both hands, DualHand (DH-) is used to specify that the tens step involved the dominant hand and the digits step involved the non-dominant hand.

Hence the four presented techniques can be described as follows: **MidAir<sup>2</sup>** is the full mid-air technique; **Touch<sup>2</sup>** is based on direct touch only; **Touch  $\times$  MidAir** combines touch for the tens and mid-air input for the units; and **DH-Touch  $\times$  MidAir** is similar to the previous one but with a bimanual approach (touch with the dominant hand, mid-air with the non-dominant). We detail the design of the baseline (mid-air only) and our three referent-based techniques below.

## 5.2 MidAir<sup>2</sup>: mid-air baseline technique

We designed a baseline using only mid-air gestures which we named MidAir<sup>2</sup> (i.e. MidAir × MidAir, using mid-air gestures for both the tens and units steps). We chose this technique as a baseline because it is one of the most commonly available input modes on HMDs (e.g. HoloLens).

When first designing the full mid-air interaction to select data on a STC, we considered raycasting with the hand or gaze to select data, followed by validation with a hand gesture (such as the AirTap on HoloLens). However, given the complexity of the information space to visualize (e.g. large dataset on a STC) and the number of task modes, this solution led to selection ambiguities between the different axes and faces. For instance, to point on the YZ face of the STC, the user must cast a ray which will also intersect the XY face. The same goes for 3D extraction, which is difficult to achieve using gaze in dense data spaces, as demonstrated in previous works [59][103][104]. One solution could be to place virtual anchors on the STC axes and faces as the primary target of the ray casting. However, this would add visual clutter to the already complex STC, introducing a bias in our study as it would be the only technique using them.

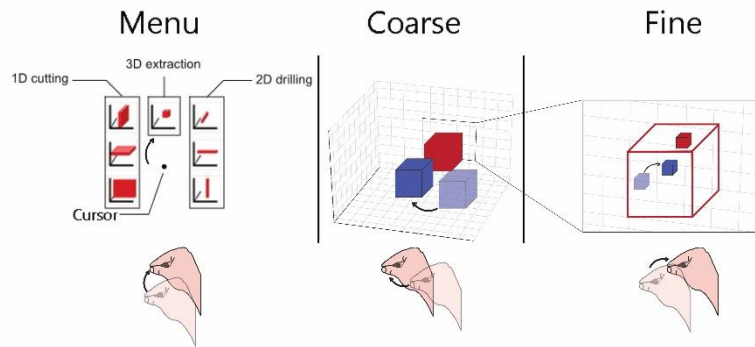


Figure 5. MidAir<sup>2</sup> interaction technique started with a radial menu used for task selection (left) before performing the data selection task (right). The cursor is displayed in blue and the target in red.

Instead, we decided to use a radial menu controlled by hand gestures. We used a radial menu to ensure that every item was equidistant from the initial cursor position, thus limiting a potential time difference between each task and the menu item selection process. The user can select the task (1D cutting, 2D drilling, 3D extraction) and the axis/plane on the radial menu, displayed next to the STC not to hinder the data visualization. To select a menu item, the user presses a handled clicker and moves the hand in two dimensions to move a 2D cursor (see Figure 5 – left). Releasing the clicker validates the item selection. There was no reset mechanism, as users could not start the selection of the tens without validating the correct menu item beforehand. We chose to rely on a physical clicker as previous work has shown that mid-air validation gestures (e.g. AirTap) are less efficient than physical clickers [88]. After this task selection on the menu, the user starts data selection on the STC by pressing the clicker again. As long as the user press the clicker, the spatial position of the hand is mapped to the virtual cursor position within the STC (see Figure 5 – right). We used a dwell time to validate the target selection. The control-display gain was set empirically through preliminary tests and ended up with a mid-air gesture of 5cm that corresponds to 1 unit in the STC. This gain was defined by asking users to perform the gesture with maximum comfort and ease of interaction. It appears to be a good compromise between a smaller CD gain that would trigger more ample gestures and potential fatigue and a higher CD-gain that would require highly accurate mid-air gestures.

To sum up, the MidAir<sup>2</sup> interaction is carried out as follows: selecting the task in the menu, moving the virtual cursor to select the tens step target, then after the dwell time validation, moving the virtual cursor to select the units step target, which ends with another dwell time validation.

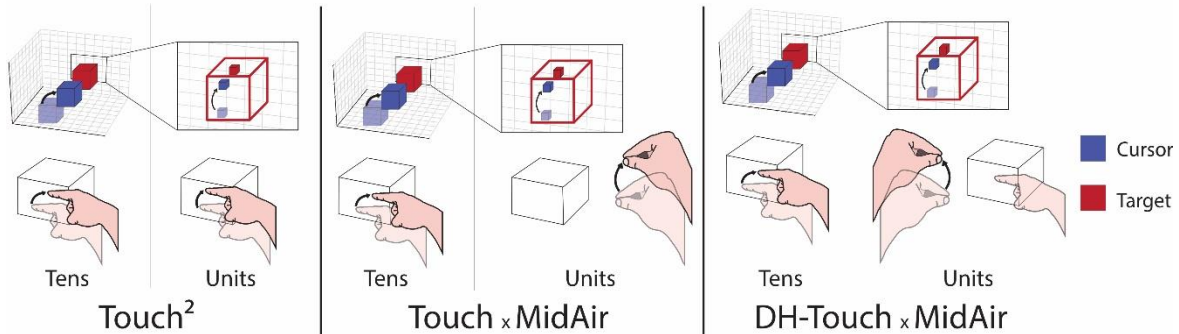


Figure 6. We designed three interaction techniques involving touch on the physical referent: Touch<sup>2</sup>, Touch x MidAir, and DH-Touch x MidAir. The cursor is displayed in blue and the target in red.

### 5.3 Touch<sup>2</sup>

In the Touch<sup>2</sup> technique, direct touch interaction on the edges or faces of the physical referent (see Figure 6 – Left) allows to perform the tens and units step selections. Each physical feature of the referent (edge, face) is associated with an axis or plane of the STC: the length of a physical referent edge/face is directly mapped to the length of the STC axis/plane (10 in our case). We used the results of the preliminary study to select the specific mappings between the physical features of the referent and the dimensions of the STC (see Figure 4), as detailed below:

- For *1D cutting*, the left and top edges of the front face of the physical referent are mapped to the X and Y axes of the STC, respectively, and the left edge of the top face to the Z axis of the STC.
- For *2D drilling*, the front, right and top faces are used to perform a drill on the XY, YZ and XZ planes of the STC, respectively.
- For *3D extraction*, the front and top faces are used sequentially to select the XY and Z values of the 3D cursor, respectively.

We also added a reset mechanism: to reset the current selection, the user can touch the back of the physical referent. Finally, to differentiate the use of the front face for the 2D drilling task from the 3D extraction task, we implemented the following state machine: when touching the front face, users start with the 2D drilling task; whenever the user touches the top face, it activates the 3D extraction task. If the user initially touches the front face by mistake, the user has to use the reset mechanism, i.e. touch the back of the physical referent to cancel the current selection.

### 5.4 Touch x MidAir

This technique resulted from the feedback of the preliminary study: some users wanted to combine touch and mid-air interaction with the same hand. Their motivation was that the surface of the physical referent might be too small to allow precise touch interaction. In this context, using the mid-air space around the physical referent makes it possible to extend the interaction space.

The Touch x MidAir technique first consists in performing the tens step using touch on the physical referent, with the same physical mappings as in the Touch<sup>2</sup> technique (see section 5.3). The reset mechanism was identical to the



Touch<sup>2</sup> technique. Then the user performs the units step in mid-air with the same hand (see Figure 6 – Center) and in the same way as the units step of the MidAir<sup>2</sup> technique (see section 5.2).

### **5.5 DH-Touch x MidAir**

The DH-Touch x MidAir technique is a bimanual variant of the previous *Touch x MidAir* technique. Instead of using the same hand for both the tens and units steps, this technique relies on touching the physical referent with the dominant hand (DH) for the tens step, followed by mid-air interaction with the non-dominant hand (NDH) for the units step (see Figure 6 – Right) as suggested by the results of the preliminary study (see section 4.2). Although the input space for the tens step requires finer movements (touching the 10 cm wide physical referent) than the input space for the units step (performing a mid-air gesture with a 50 cm amplitude, see section 5.2), this solution was the most often proposed one in the preliminary study (5 participants out of 12).

## **6 STUDY 1: DATA CUTTING, DRILLING AND EXTRACTION**

This study aimed to compare our three referent-based techniques involving mid-air or touch interaction on the physical referent, with the baseline technique relying solely on mid-air interaction, for the three data analysis tasks (1D cutting, 2D drilling and 3D extraction).

### **6.1 Task**

As described in section 5.1, selecting the values from 1 to 100 is split into two steps: a tens step to define the tens digit and a units step to define the unit's digit. We uniformly distributed the tens and units target positions along the STC axes, planes and 3D space. For both the tens and units steps, users had to move a virtual cursor (blue coloured, see Figure 7) and place it on a target displayed on the STC (red coloured, see Figure 7). The cursor shape differed according to the task: a plane for 1D cutting, a ray for 2D drilling or a cube for 3D extraction.

Participants were sitting in front of a table with the physical referent. The general instruction was to complete the selection tasks quickly and precisely. Participants completed each trial as follows: 1) on startup, the instructions for the subsequent trial were displayed using the HMD and participants had to place their hands at rest on the table. 2) To begin the trial, the user had to press a handled clicker with their dominant hand. 3) The STC was displayed with the tens step target. 4) Once the user placed the cursor on the target, it turned green after 500ms, validating the tens step selection. 5) The same process applied for the units step. For the baseline, the protocol included an additional step (before 3): to select the desired task on the radial menu using the handled clicker and mid-air movement of the dominant hand.

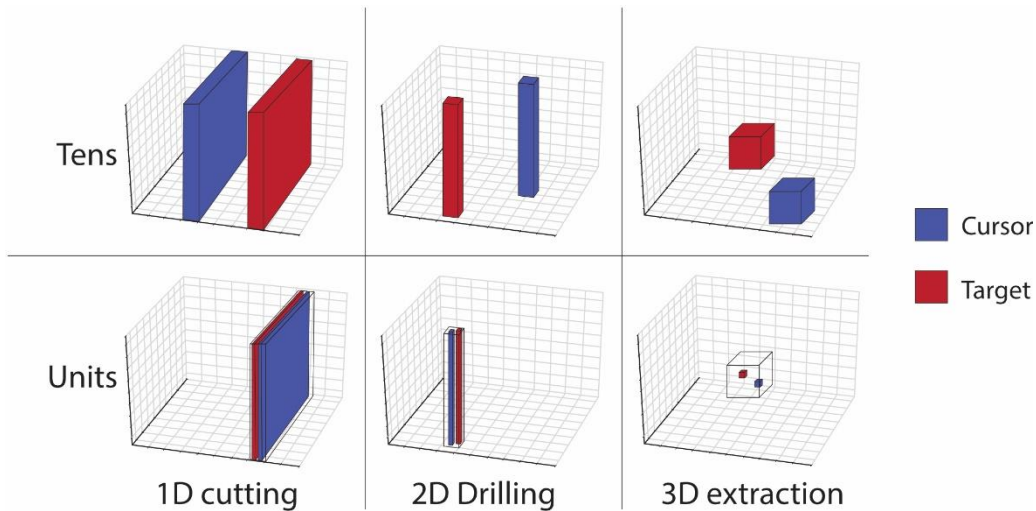


Figure 7. Pointing tasks in the space-time cube in tens and units steps: X axis for 1D cutting, XZ face for 2D drilling and 3D extraction. The cursor is displayed in blue and the target in red.

## 6.2 Apparatus and implementation

The STC was displayed above the physical model, in front of the user, 1m away from him, using HoloLens v1. The HoloLens resolution was 1268x720 with a FoV of 30°x17.5° for each eye. The physical referent consisted of a cubic-shaped box fixed on a static horizontal surface made of LEGO® bricks (60cm from the user). The LEGO® surface was glued on the horizontal surface. The dimensions of the box representing the physical referent were 10cm x 10cm x 10cm, similar to cubic-shaped smart devices (Figure 2).

To detect the touch interaction on the physical referent, we used an OptiTrack infrared tracking system [79]. The calibration procedure of the system provided a resulting precision of 1.6mm and 1.4° on average for the tracking. Users had to wear gloves equipped with infrared markers. We also placed markers on the physical referent, ensuring that none of these markers interfered with the interaction gestures (see Figure 8). The system tracked the hand position in 1m x 1m x 1m in front of the users. To remove any technological bias between techniques and to overcome the limited mid-air gestures tracking space of HoloLens, we also tracked the hand in mid-air using the infrared system. We used a mini-mouse attached to the palm as a clicker (Figure 8 – Center-Right). We disinfected the HoloLens between sessions, and we ensured participants respected social distancing and barrier gestures to prevent any risks related to COVID-19.

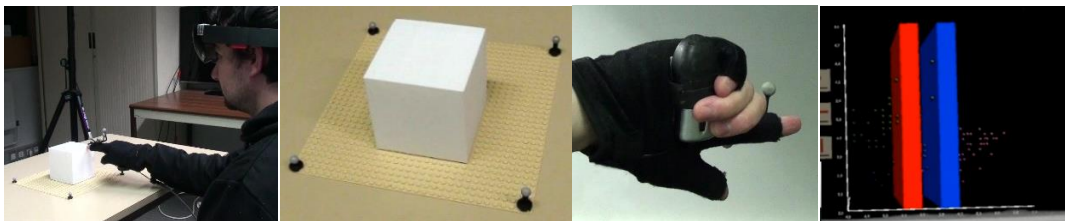


Figure 8. From left to right: The user with a HoloLens touches the physical referent in our experimental setup. Infrared markers were placed on the physical referent and the user's hands. First-person view of the STC rendering with the cursor in blue and the target in red.

### 6.3 Participants

We recruited 12 participants at the local university, all with a background in computing science (26.0 years old on average, 3 females, 9 males). None of them received compensation for their participation. Six of them participated in the preliminary study and 8 had previous experience with AR devices. The analysis of the experimental data did not reveal any difference between the participants who had taken part in the preliminary study and those who did not (see Appendices, Figure 18 and Table 1). Hence, we analyzed the results of both groups together [66].

### 6.4 Experimental design and collected data

The experiment followed a 4x3 within-subjects design, with Interaction Technique (MidAir<sup>2</sup>, Touch<sup>2</sup>, Touch x MidAir, DH-Touch x MidAir) and Task (1D cutting, 2D drilling, 3D extraction) as factors. The study was divided into four blocks, each corresponding to one technique, and counterbalanced across participants. For each interaction technique, participants had to perform 9 training trials in the same order: 3 trials for 1D cutting, 3 trials for 2D drilling and 3 trials for 3D extraction. They were followed by 18 trials (2 repetitions of 1D drilling on X, Y and Z axes + 2 repetitions of 2D cutting on XY, XZ and YZ planes and 6 repetitions of 3D extraction). For the experiment, we collected 12 participants x 4 interaction techniques x 18 trials = 864 trials.

### 6.5 Collected data and data analysis

We measured the trial completion time and the number of errors (bad menu selections for MidAir<sup>2</sup>, bad edge/face selections of the physical referent for touch interaction techniques). After each interaction technique, users had to fill out a Borg scale [16] and a SUS questionnaire [44]. At the end of the experiment, users had to rank the interaction techniques in terms of preference.

In this work, we relied on estimation techniques with 95% confidence intervals and ratio analysis as recommended by the APA [85]. Furthermore, there is a push in the CHI community to shift focus towards effect sizes instead of p-values (<https://www.aviz.fr/badstats#sece>). Using CIs rather than p-values allows for a more nuanced analysis of the direction and magnitude of the effect instead of a dichotomous inference based on p-values. The ratio is an intra-subject measurement that expresses the effect size (pair-wise comparison) and is computed between each geometric means. All CIs are 95% BCa bootstrap confidence interval. For the reader more used to interpret the p-values, a parallel might be drawn with results obtained through the estimation technique and confidence interval reports (see Figure 3 in [57]). Scripts used to compute the geometric mean and confidence intervals were used in [119] and are available online ([www.aviz.fr/ci](http://www.aviz.fr/ci)).

## 6.6 Results

### 6.6.1 Errors

We could not find any difference among the four interaction techniques regarding errors. Users rarely selected the wrong menu item in the MidAir<sup>2</sup> technique (1.3%) and seldom interacted on the wrong edge/face of the physical referent (1.2%).

### 6.6.2 Completion time

We first analyze the results for each task independently (1D cutting, 2D drilling and 3D extraction).

**Task 1: 1D cutting.** For 1D cutting (see Figure 9, top-left), we found that when using *MidAir*<sup>2</sup> (7.93s, CI: [7.06, 9.25]) participants took more time than with the *Touch*<sup>2</sup> (4.19s, CI: [3.85, 4.49]), *Touch* × *MidAir* (5.17s, CI: [4.43, 5.98]) and *DH-Touch* × *MidAir* (5.10s, CI: [4.56, 6.44]).

An intra-subject analysis based on the ratio (see Figure 9, bottom-left) confirms these results. It establishes that when using the *MidAir*<sup>2</sup> technique, participants take 88% (1.88, CI: [1.75, 2.08]) more time than with *Touch*<sup>2</sup>, 58% (1.58, CI: [1.39, 1.75]) more time than with *Touch* × *MidAir* and 61% (1.61, CI: [1.42, 1.90]) more time than with *DH-Touch* × *MidAir*.

Our results reveal that the use of a physical referent facilitates data selection. Indeed, touching the physical referent on its edge offers physical boundaries for the gesture, allowing to place the cursor faster on a value close to one boundary. For instance, a user could place the index finger on the left side of the edge to select a value lower than five, reducing the displacement of the finger on the edge. This type of shortcut is not possible for mid-air interaction.

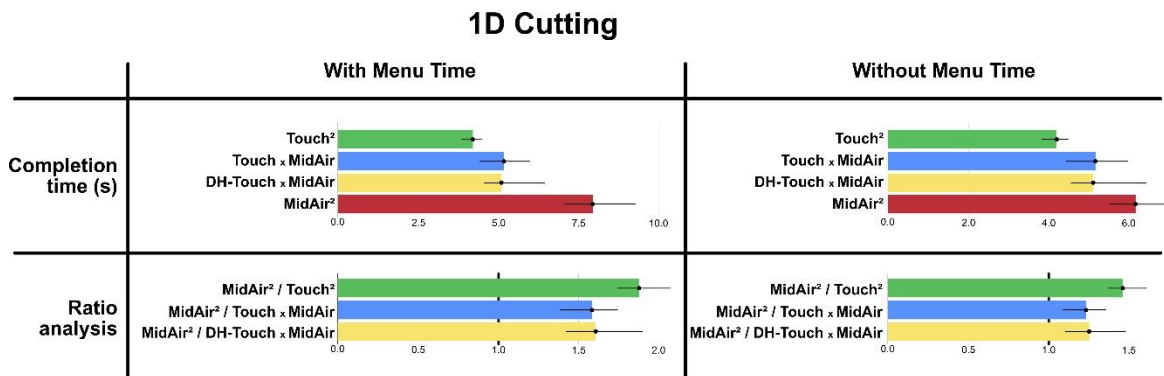


Figure 9. Average completion time (in seconds) for 1D cutting (top-left) and completion time ratio (bottom-left) with menu time and without menu time (respectively top-right and bottom right) (95% CIs).

As described in section 5.2, when using *MidAir*<sup>2</sup>, users must first select the task in a radial menu. This design choice requires users to spend time on the radial menu to define the selection task: the menu time. Although this was the best design we could think at, in case another design could be identified that would allow getting rid of the radial menu, we wanted to analyze our results without this menu time. Even so, it appears that *Touch*<sup>2</sup> still outperform *MidAir*<sup>2</sup> and that *Touch* × *MidAir* and *DH-Touch* × *MidAir* still require less time than *MidAir*<sup>2</sup>, but the difference is lessened (see Figure 9 – Right). The means, CIs and intra-subject ratio can be found in the Appendices, Table 3 and Table 4.

**Task 2: 2D Drilling.** For 2D drilling (see Figure 10, top-left), participants also took more time with the *MidAir*<sup>2</sup> (10.63s, CI: [9.08, 12.83]) technique than with *Touch*<sup>2</sup> (8.07s, CI: [6.97, 9.53]) or with *Touch* × *MidAir* (8.10s, CI: [6.89, 9.27]). There is no clear difference between *MidAir*<sup>2</sup> and *DH-Touch* × *MidAir* (9.34s, CI: [7.53, 14.62]).

A ratio analysis confirms the two first results of 2D drilling (see Figure 10, bottom-left): *MidAir*<sup>2</sup> requires 35% (1.35, CI: [1.18, 1.58]) more time than *Touch*<sup>2</sup> and 33% (1.33, CI: [1.18, 1.49]) more time than *Touch* × *MidAir*. However, unlike the average completion time analysis, the intra-subject ratio analysis reveals that *MidAir*<sup>2</sup> requires 22% (1.22, CI: [1.09, 1.39]) more time than *DH-Touch* × *MidAir*.

As for 1D cutting, the faces of the physical referent offer tangible cues (physical delimiters) for the interaction space, which explains our results. It allows the user to place the initial cursor position closer to the desired value and then

refine the position using small movements. Using mid-air interaction, users have no choice but to start at the centre of the virtual space and perform the whole movement to the target position.

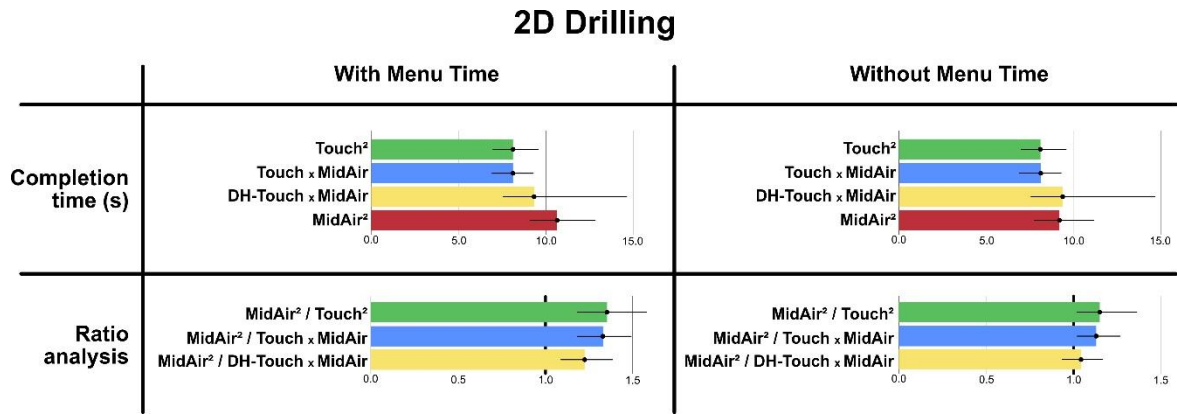


Figure 10. Average completion time (in seconds) for 2D drilling (top-left) and completion time ratio (bottom-left) with menu time and without menu time (respectively top-right and bottom right) (95% CIs).

The analysis without the menu time also establishes that *Touch<sup>2</sup>* and *Touch x MidAir* require less time than *MidAir<sup>2</sup>*. However, we found no difference between *MidAir<sup>2</sup>* and *DH-Touch x MidAir* (see Figure 10 – Right). The means, CIs and intra-subject ratio can be found in the Appendices, Table 5 and Table 6.

**Task 3: 3D extraction.** Results differ for 3D extraction (see Figure 11 – top): participants took more time with *Touch<sup>2</sup>* (17.4s, CI: [14.81, 21.44]) than with the three other techniques, namely *Touch x MidAir* (12.8s, CI: [10.92, 15.21]); *DH-Touch x MidAir* (13.6s, CI: [11.70, 17.09]) and *MidAir<sup>2</sup>* (11.8s, CI: [10.57, 13.77]). We found no difference between the three last techniques.

The intra-subject ratio analysis (see Figure 11 – bottom) establishes in addition that the *MidAir<sup>2</sup>* requires 28% (0.72, CI: [0.63, 0.84]) less time than *Touch<sup>2</sup>* and 11% (0.89, CI: [0.85, 0.93]) less time than *DH-Touch x MidAir*. The ratio analysis does still not reveal any clear difference between *MidAir<sup>2</sup>* and *Touch x MidAir*.

For 3D extraction, touch interaction requires two successive actions to select the tens (as well as for the units). Indeed, to select the tens (respectively units), a first touch on the front face selects a value on the XY plane, followed by a second touch on the top face to select the value on the Z axis. These two actions performed sequentially require more time, inducing longer absolute times for the three techniques involving such interaction on the physical referent. Meanwhile, selecting the tens with *MidAir<sup>2</sup>* only requires a single 3D gesture, as the hand's position in mid-air is directly mapped to one target inside the 3D STC. With touch-based techniques, beyond having to perform twice as many actions as with mid-air techniques, participants also have to alternatively touch different faces of the physical referent, which ends up making touch-based techniques slower than the others for 3D extraction.

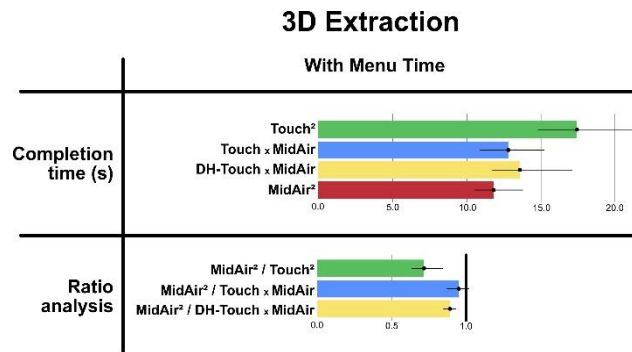


Figure 11. Average completion time (in seconds) for 3D extraction (top) and completion time ratio (bottom) (95% CIs).

As opposed to the two other tasks, we do not refine this analysis by detailing the menu time of *MidAir<sup>2</sup>* as our techniques are slower than it.

### 6.6.3 Fatigue

Results clearly show that *MidAir<sup>2</sup>* induces more fatigue than *Touch<sup>2</sup>* and *DH-Touch x MidAir* techniques: the average scores on the 10-values Borg scale are respectively 5.0 CI: [4.31, 5.62], 2.0 CI: [1.38, 2.31] and 2.7 CI: [2.0, 3.38]. We observed only a slight difference between *MidAir<sup>2</sup>* and *Touch x MidAir* (3.84 CI: [3.23, 4.38]).

Based on the informal feedback from participants, these differences can be explained by touch interaction on the physical referent (*Touch<sup>2</sup>*) being less tiring than mid-air interaction. For *DH-Touch x MidAir*, participants expressed that the fatigue is mainly related to the units step (mid-air interaction), which explains why this technique was slightly less tiring than *MidAir<sup>2</sup>*.

### 6.6.4 System Usability Score & preference

Results only revealed only minor differences among the SUS [44] scores for the *MidAir<sup>2</sup>*, *Touch<sup>2</sup>*, *Touch x MidAir* and *DH-Touch x MidAir* techniques (respectively 82 CI: [76.34, 85.96], 88 CI: [77.23, 93.85], 86 CI: [74.49, 92.12] and 83 CI: [76.41, 88.46] average scores). The score obtained for all techniques corresponds to a “good” or “excellent” usability.

Regarding the ranking of the interaction techniques, we found that *Touch<sup>2</sup>* and *Touch x MidAir* are most often ranked in first and second places (9 and 7 times, respectively). *MidAir<sup>2</sup>* is most often ranked in second and third place (7 times, respectively) and *DH-Touch x MidAir* is most often ranked as the least preferred (11 times respectively).

Interestingly, although we did not measure it, we noticed that participants did not often look at the physical referent, keeping their attention on the STC instead. Hence, beyond the performance, fatigue and usability advantages, touch-based input on the physical referent allows for eyes-free interaction, which lets the user focus on the data visualization itself.

### 6.6.5 Summary

To sum up, we found that participants took less time to perform 1D cutting and 2D drilling with *Touch<sup>2</sup>* and *Touch x MidAir* than with *MidAir<sup>2</sup>*. The time gain varied from 58% to 88% for 1D Cutting and 22% to 35% for 2D drilling. These techniques induced less fatigue, had better SUS scores than *MidAir<sup>2</sup>* and were the preferred techniques.

We also found that for 3D extraction, users needed 11% to 28% less time using *MidAir*<sup>2</sup> than *Touch*<sup>2</sup> and *Touch* × *MidAir*. Hence, we decided to re-design the touch interaction for 3D extraction to remove this gesture redundancy.

## 7 STUDY 2: TOUCH TECHNIQUES FOR 3D EXTRACTION

This study aims to evaluate the design of new touch interaction techniques for efficient 3D Extraction, compared to the *MidAir*<sup>2</sup> baseline.

### 7.1 Interaction techniques

The results and subjective feedback collected in the previous study indicate that the main problem of performing the 3D extraction task with touch input was the need to sequentially touch two faces of the physical referent to select the tens and again to select the units. Indeed, to select a 3D point, a first touch selects the values on the XY plane and a second touch selects a value along the Z dimension. This breaks the interaction flow and hinders the performance of touch techniques.

To answer this limitation, instead of using two sequential touch actions, we designed two interaction techniques based on the simultaneous use of both hands. We named those techniques using the same naming convention described in section 5.1, i.e. as a product of the input mode (IM) used for each selection step:  $\langle IM\_tens \times IM\_units \rangle$ . We extended this convention with the prefix Bi- (Bimanual) to describe a two-hand interaction technique in which both hands were used during each step of the selection process. As a result, the two new techniques are named *Bi-Touch*<sup>2</sup> and *Bi-TouchMidAir*<sup>2</sup>.

*Bi-Touch*<sup>2</sup> is based on touch input only. It consists in using the dominant hand on the front side of the physical referent to perform XY movements while simultaneously using the non-dominant hand on the left side (for a right-handed user) of the physical referent for the Z axis. The two hands can be placed simultaneously on the physical referent, allowing parallel navigation along the three dimensions while removing the need to move the finger sequentially from one face of the physical referent to another.

*Bi-TouchMidAir*<sup>2</sup> is based on the combined use of touch and mid-air interaction. It simultaneously uses the dominant hand on the front side of the physical referent to perform XY movements and the non-dominant hand in mid-air, moving forward/backwards to adjust the Z axis. This combination defines the tens and again defines the units.

The *MidAir*<sup>2</sup> technique, which we use again as a baseline, was precisely the same as in the previous study.

### 7.2 Task and apparatus

Participants had to perform the 3D extraction task only; the apparatus was identical to study 1. The only difference with study 1 regarding the implementation relies on the feedback for 3D extraction. Informal comments on the previous study revealed that some participants would like feedback to facilitate alignment perception. To this end, we added a line orthogonal to the XY face of the STC and passing through the cursor (blue). As such, when the line intersects the target (red), it signifies that both are aligned and the final adjustment only requires moving along the Z axis. As we didn't know if the feedback would impact the results, we tested both with and without such feedback, counterbalancing the condition across participants.

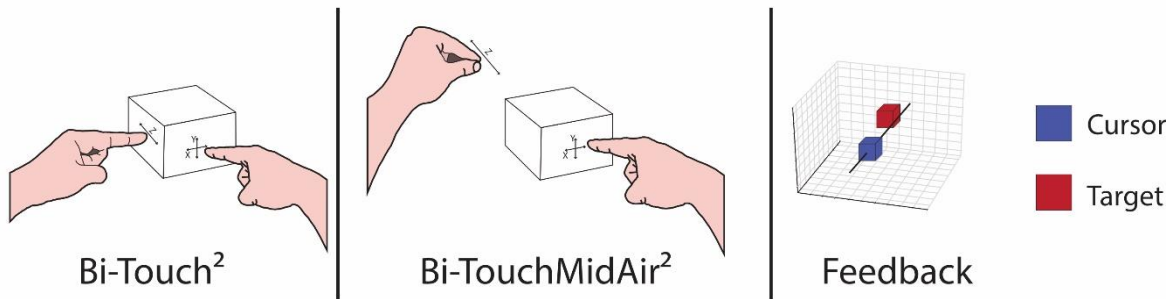


Figure 12. We designed two bimanual Interaction techniques for 3D extraction: *Bi-Touch<sup>2</sup>* (left) and *Bi-TouchMidAir<sup>2</sup>* (Center). For both, the tens and units steps are performed similarly. We also investigated adding depth perception feedback (Right).

### 7.3 Participants, experimental design and collected data

We recruited 12 participants for the experiment, all with a background in computing science (26.6 years old on average, 4 females, 8 males). None of them received compensation for their participation. Six of them took part in study 1. The analysis of the experimental data did not reveal any difference between the participants who had taken part in the previous study and those who did not (see Appendices, Figure 19 and Table 2). Hence, we analyzed the results of both groups together [66].

The experiment followed a 3x2 within-subjects design, with the Interaction technique (*MidAir<sup>2</sup>*, *Bi-Touch<sup>2</sup>*, *Bi-TouchMidAir<sup>2</sup>*) and Depth feedback (with or without) as factors. The study was divided into three blocks, each corresponding to one technique, which were counterbalanced across participants. For each interaction technique, users had to perform 8 training trials (4 with the depth feedback, 4 without), followed by 24 trials (12 with the feedback and 12 without). We uniformly distributed the tens' and units' target positions along the 3D space differently for all participants. In total, we collected 12 participants x 4 interaction techniques x 2 feedback x 12 trials = 864 trials.

We measured the trial completion time and the number of errors (wrong menu selections for *MidAir<sup>2</sup>* and input on the wrong face of the physical referent). After each interaction technique, users had to fill out a Borg scale [16] and a SUS questionnaire [44]. At the end of the experiment, users had to rank the interaction techniques in terms of preference.

### 7.4 Results

We conducted the same data analysis as in the previous study, i.e. using estimation techniques with 95% confidence intervals and ratio analysis (see section 6.6).

#### 7.4.1 Errors

Users rarely selected the wrong menu item in the *MidAir<sup>2</sup>* technique (1.1%, we found 1.3% in Study 1). We did not record any input on the wrong faces of the physical referent during the experiment.



### 7.4.2 Completion Time

With *Bi-Touch<sup>2</sup>* (7.98 ms, CI: [7.31, 8.56]), participants clearly took less time to complete the 3D extraction task than with *Bi-TouchMidAir<sup>2</sup>* (9.69 ms, CI: [8.32, 10.27]) and *MidAir<sup>2</sup>* (11.25 ms, CI: [9.42, 12.88]), see Figure 13 – top-left. While the difference is less important, *Bi-TouchMidAir<sup>2</sup>* also requires less time than *MidAir<sup>2</sup>*.

The ratio analysis establishes that with *MidAir<sup>2</sup>*, participants took 38% more time than with *Bi-Touch<sup>2</sup>* (1.38, CI: [1.25, 1.51]) and 15% more time than with *Bi-TouchMidAir<sup>2</sup>* (1.15, CI: [1.03, 1.25]), see Figure 13 – bottom-left.

These results further confirm that using a physical referent efficiently performs 3D extraction tasks on STCs. We believe that the physical support facilitates interaction compared to mid-air, where the hand position is not tied to a pre-defined surface, thus explaining our results. First, the physical referent provides users with visual and haptic hints: the physical features of the referent (edges, corners) help to anticipate the minimal, maximal and current position on the axis/face. Second, the physical referent may contribute to a more stable pointing gesture.

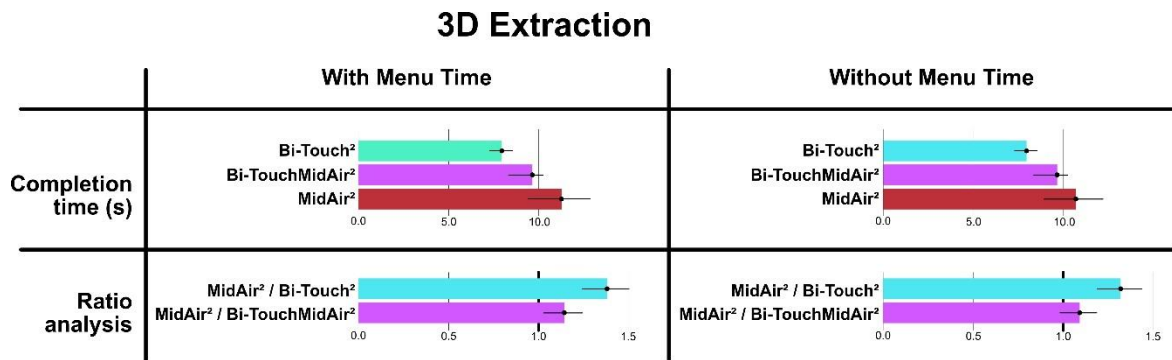


Figure 13. Completion time in seconds (top-left) and the corresponding ratio (bottom-left) for the 3D extraction task with menu time and without menu time (respectively top-right and bottom-right) (95% CIs).

Similarly to the previous study, we further analyze these results when subtracting the menu time, i.e. the time to select a menu item with the *MidAir<sup>2</sup>* technique. The results confirm that *Bi-Touch<sup>2</sup>* and *Bi-TouchMidAir<sup>2</sup>* still require less time than *MidAir<sup>2</sup>* (see Figure 13 – Right). The means, CIs and intra-subject ratio can be found in the Appendices, Table 7 and Table 8.

### 7.4.3 Depth Feedback

While all users agreed on the usefulness of the cursor feedback to improve depth perception, we found no statistical difference between the trials performed with and without feedback across all interaction techniques.

### 7.4.4 Fatigue

Results show a clear difference in terms of fatigue between the *MidAir<sup>2</sup>* on one hand and the *Bi-Touch<sup>2</sup>* and *Bi-TouchMidAir<sup>2</sup>* techniques on the other hand (respectively 4.5 CI: [3.8, 4.9], 1.9 CI: [1.4, 2.3] and 2.7 CI: [2.3, 3.1] average scores on the Borg scale).

Informal feedback confirmed these results: “Mid-air interaction induces more fatigue than the two others (P2)” or “I won’t be able to select targets all day with this technique, or I’ll have to take breaks (P3)”.

#### 7.4.5 System Usability Score (SUS) & preference

We found only a slight difference between the SUS scores of the *MidAir<sup>2</sup>*, *Bi-TouchMidAir<sup>2</sup>* and *Bi-Touch<sup>2</sup>* techniques. All of them offered “good” or “excellent” usability (respectively 81 CI: [76.5, 84.4], 79 CI: [71.7, 82.9] and 88 CI: [86.7, 91.5] average scores). Informal feedback revealed that this difference is due to mid-air interaction being less appreciated than touch interaction.

Regarding the ranking of the interaction techniques, *Bi-Touch<sup>2</sup>* was most often ranked in the first place (11 out of 12 times), *MidAir<sup>2</sup>* was most often ranked in second place (8 times), and *Bi-TouchMidAir<sup>2</sup>* most often ranked last (9 times respectively).

#### 7.4.6 Summary

Overall, we found that when using *Bi-Touch<sup>2</sup>*, participants performed the 3D extraction task faster and with less fatigue and that participants preferred (SUS and ranking) this technique over the *MidAir<sup>2</sup>* baseline.

### 8 STUDY 3: COMPARING VERTICALLY AND HORIZONTALLY FIXED PHYSICAL REFERENTS

In our initial studies, we considered only seated participants. However, as illustrated in the different scenarios presented in section 3.1, some involve users standing in front of a vertically fixed physical referent (e.g. a sensor fixed on the wall). Therefore, to confirm the validity of our results, we conducted a third study in which the user is standing in front of either a physical referent laid on a horizontal surface (Figure 14 – Left) or attached to a vertical wall (Figure 14 – Right).



Figure 14. A user standing in front of a horizontal physical referent (Left) or vertical physical referent (Right).

#### 8.1 Interaction techniques

In this experiment, we compared the mid-air interaction (*MidAir<sup>2</sup>*, baseline) to the more efficient and preferred direct touch interaction techniques for 1D cutting, 2D drilling and 3D extraction from our previous studies. We thus used *Touch<sup>2</sup>* (see section 5.3) for 1D cutting and 2D extraction, and *Bi-Touch<sup>2</sup>* (see section 7.1) for 3D extraction.

#### 8.2 Task and apparatus

Participants had to perform the seven tasks described in section 6.1: 1D cutting for X, Y and Z axes, 2D drilling for XY, XZ and YZ planes and 3D extraction. The physical referent, interaction space and HoloLens device were identical to the previous studies. The user is standing in front of the referent (distance of 60cm), which can either be horizontally glued on a table or vertically fixed on a wall. The physical referent was always placed at the height of the user's elbow to limit biases due to the height difference between the user's hand and the physical referent in

both conditions. Compared to study 2, we removed the visual feedback for the 3D extraction task as it had no effect on results.

### 8.3 Participants, experimental design and collected data

We recruited 12 participants for the experiment, all with a background in computing science (24.6 years old on average, 5 females, 7 males). None of them received compensation for their participation. None of them took part in any of the previous studies.

The experiment followed a 2x2 within-subjects design, with the Interaction technique (*MidAir*<sup>2</sup> for 1D, 2D and 3D tasks, versus *Touch*<sup>2</sup> for 1D and 2D tasks and *Bi-Touch*<sup>2</sup> for the 3D task) and Physical referent position (Vertically fixed or Horizontally fixed) as factors. We divided the study into four blocks, each corresponding to the combination of one interaction technique and one physical referent position. The order of the blocks was counterbalanced across participants. For each interaction technique, participants had to perform 9 training trials in the same order: 3 trials for 1D cutting, 3 trials for 2D drilling and 3 trials for 3D extraction. They were followed by 18 trials (2 repetitions of 1D drilling on X, Y and Z axes + 2 repetitions of 2D cutting on XY, XZ and YZ planes and 6 repetitions of 3D extraction). We uniformly distributed the tens' and units' target positions differently for all participants. For the experiment, we collected 12 participants x 2 interaction techniques x 2 physical referent positions x 18 trials = 864 trials.

We measured the trial completion time and the number of errors (bad menu selections for *MidAir*<sup>2</sup> and input on the wrong face of the physical referent for the direct touch techniques). After each interaction technique, users had to fill a Borg scale [16] and a SUS questionnaire [44]. At the end of the experiment, users had to rank the interaction techniques in terms of preference.

### 8.4 Results

We conducted the same data analysis as in the previous studies, i.e. using estimation techniques with 95% confidence intervals and ratio analysis (see section 6.6).

#### 8.4.1 Errors

Users rarely selected the wrong menu item in the *MidAir*<sup>2</sup> technique (1.1% versus 1.3% in Study 1 and 1.1% in Study 2). They also seldom interacted on the wrong edge/face of the physical referent (1.0% versus 1.2% in Study 1 and 0% in study 2).

#### 8.4.2 Completion Time

We first analyze the results for each task independently (1D cutting, 2D drilling and 3D extraction).

**Task 1: 1D cutting.** We found no difference between horizontal (4.34s, CI: [3.98, 4.65]) and vertical (4.61s, CI: [3.94, 5.41]) physical referents when participants used *Touch*<sup>2</sup> (see Figure 15, top-left). We also found no difference between horizontal (7.58s, CI: [6.68, 8.99]) and vertical (7.43s, CI: [6.60, 9.24]) physical referents when participants used *MidAir*<sup>2</sup>. These completion times are in line with our previous findings when the user was seated in front of a horizontally fixed physical referent (section 6.6.2, *Touch*<sup>2</sup> required 4.19s and *MidAir*<sup>2</sup> required 7.93s): they show that performing a 1D cutting with *MidAir*<sup>2</sup> requires more time than with *Touch*<sup>2</sup>.

An intra-subject analysis based on the ratio (see Figure 15, bottom-left) confirms these results. It establishes that when using the *MidAir*<sup>2</sup> technique, participants take 73% (1.73, CI: [1.60, 1.92]) more time than with *Touch*<sup>2</sup> with a horizontal physical referent and 66% (1.66, CI: [1.47, 1.87]) more time with a vertical physical referent.

As such, the results of this study (user standing), compared to our previous studies' findings (user seated), indicate that the user's position (standing versus seated) doesn't seem to interfere with the 1D cutting time.

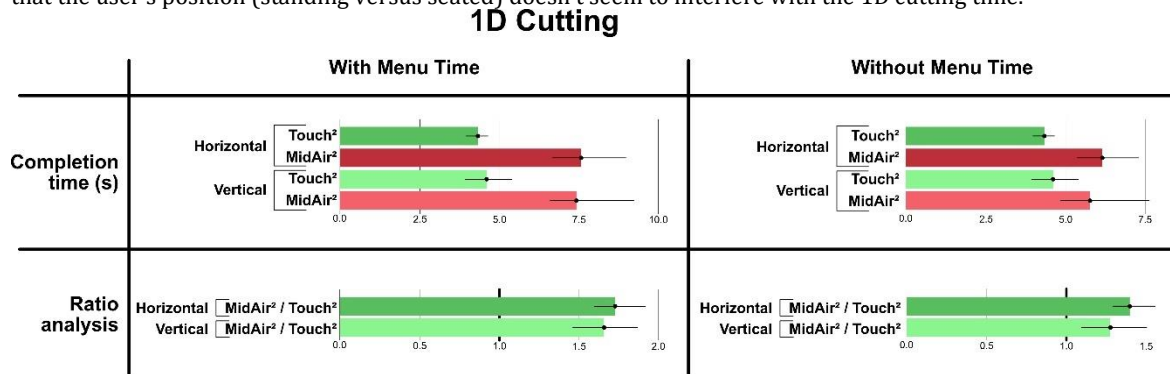


Figure 15. Average completion time (in seconds) for 1D cutting (top-left) and completion time ratio (bottom-left) with menu time and without menu time (respectively top-right and bottom-right) (95% CIs).

Similarly to the previous studies, we further analyze our results when subtracting the time to select a menu item with the *MidAir*<sup>2</sup> technique. The results confirm that *Touch*<sup>2</sup> requires less time than *MidAir*<sup>2</sup> when interacting with a horizontal or vertical physical referent (see Figure 15 – Right). The means, CIs and intra-subject ratio can be found in Appendices, Table 9 and Table 10.

**Task 2: 2D drilling.** We found no difference between horizontal (7.29s, CI: [6.31, 8.20]) and vertical (7.59s, CI: [6.42, 8.76]) physical referents when participants used *Touch*<sup>2</sup> (see Figure 16, top-left). We also found no difference between horizontal (10.26s, CI: [8.69, 12.27]) and vertical (10.28s, CI: [8.23, 15.61]) physical referents when participants used *MidAir*<sup>2</sup>. As for 1D cutting, these completion times are in line with our previous findings when the user was seated in front of a horizontally fixed physical referent (section 6.6.2, *Touch*<sup>2</sup> required 8.07s and *MidAir*<sup>2</sup> required 10.63s): they show that performing a 2D drilling with *MidAir*<sup>2</sup> requires more time than with *Touch*<sup>2</sup>.

An intra-subject analysis based on the ratio (see Figure 16, bottom-left) confirms these results. It establishes that when using the *MidAir*<sup>2</sup> technique, participants take 41% (1.41, CI: [1.28, 1.66]) more time than with *Touch*<sup>2</sup> with a horizontal physical referent, and 35% (1.35, CI: [1.16, 1.83]) more time with a vertical physical referent.

Consequently, comparing these results with our previous studies' findings (user seated) indicates that the user's position (standing versus seated) doesn't seem to interfere with the 2D drilling time.

## 2D Drilling

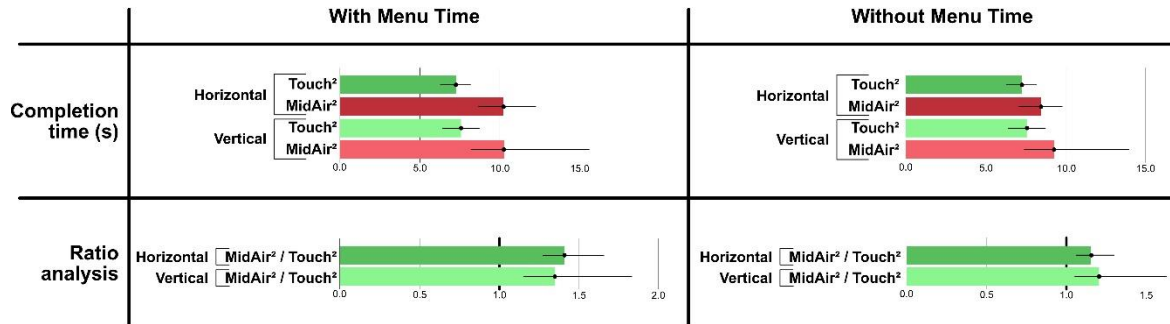


Figure 16. Average completion time (in seconds) for 2D drilling (top-left) and completion time ratio (bottom-left) with menu time and without menu time (respectively top-right and bottom-right) (95% CIs).

An analysis without the menu time again confirms that Touch<sup>2</sup> requires less time than MidAir<sup>2</sup> when interacting with a horizontal or vertical physical referent (see Figure 15 – Right). The means, CIs and intra-subject ratio can be found in Appendices, Table 11 and Table 12.

**Task 3: 3D extraction.** For this task, results reveal a difference between horizontal (9.80s, CI: [8.86, 10.45]) and vertical (8.22s, CI: [7.49, 8.96]) physical referents when participants used *Bi-Touch*<sup>2</sup> (see Figure 17, top-left). However, we found no difference between horizontal (12.32s, CI: [10.73, 14.02]) and vertical (11.96s, CI: [10.06, 13.88]) physical referents when participants used *MidAir*<sup>2</sup>. However, these completion times still establish that performing a 3D extraction with *MidAir*<sup>2</sup> requires more time than with *Bi-Touch*<sup>2</sup> when the user is standing. This is therefore in line with our previous findings established when the user was seated in front of a horizontally fixed physical referent (section 7.4.2, *Bi-Touch*<sup>2</sup> required 7.98s and *MidAir*<sup>2</sup> required 11.25s)

An intra-subject analysis based on the ratio (see Figure 17, bottom-left) confirms these results and establishes that when using the *MidAir*<sup>2</sup> technique, participants take 25% (1.25, CI: [1.15, 1.35]) more time than with *Bi-Touch*<sup>2</sup> with a horizontal physical referent, and 43% (1.43, CI: [1.30, 1.56]) more time with a vertical physical referent.

### 3D Extraction

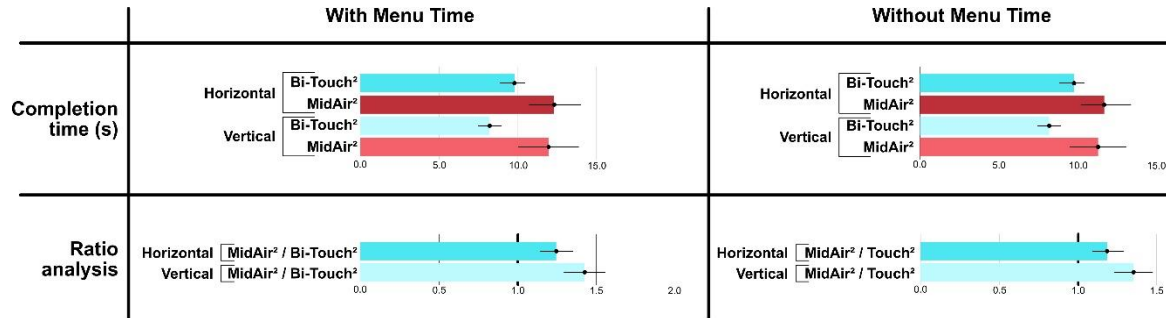


Figure 17. Average completion time (in seconds) for 3D extraction (top-left) and completion time ratio (bottom-left) with menu time and without menu time (respectively top-right and bottom-right) (95% Cis).

An analysis without the menu time confirms that Bi-Touch<sup>2</sup> requires less time than MidAir<sup>2</sup> when interacting with a horizontal or vertical physical referent (see Figure 15 – Right). The means, Cis and intra-subject ratio analysis can be found in Appendices, Table 13 and Table 14.

#### 8.4.3 Fatigue

To assess the participant's perceived exertion, we relied on the Borg Scale. We found that using a horizontal (3.08, CI: [2.67, 3.33]) referent is perceived to be more tiring than a vertical (1.5, CI: [1.0, 1.83]) physical referent when participants used direct touch techniques (*Touch<sup>2</sup>* for 1D cutting and 2D drilling; *Bi-Touch<sup>2</sup>* for 3D extraction). However, we found no difference between horizontal (5.0, CI: [4.33, 5.50]) and vertical (5.08, CI: [4.58, 5.50]) physical referents when participants used *MidAir<sup>2</sup>*. With *MidAir<sup>2</sup>*, the fatigue perceived by participants is thus higher than with our direct touch techniques, which aligns with our previous findings (sections 6.6.3 and 7.4.4).

#### 8.4.4 System Usability Score (SUS) & preference

We found no difference between horizontal (90.83, CI: [88.54, 93.13]) and vertical (92.29, CI: [90.42, 94.17]) physical referents when participants used direct touch techniques (*Touch<sup>2</sup>* for 1D cutting and 2D drilling; *Bi-Touch<sup>2</sup>* for 3D extraction). We also did not find any difference between horizontal (84.17, CI: [80.63, 86.67]) and vertical (85.0, CI: [82.08, 87.08]) physical referents when participants used *MidAir<sup>2</sup>*. These results are in line with our previous findings (sections 6.6.4 and 7.4.5), stating that participants gave a better SUS score to direct touch techniques (qualified as excellent) than to *MidAir<sup>2</sup>* (qualified as good).

Regarding the ranking of the interaction techniques, direct touch interaction with a vertical physical referent was most often ranked in the first place (8 out of 12 times), followed by direct touch interaction on a horizontal plane (4 out of 12 times). *MidAir<sup>2</sup>* was always placed at the third or fourth rank on a vertical or horizontal physical referent.

#### 8.4.5 Summary

Overall, our results confirm the findings of our previous studies: *MidAir<sup>2</sup>* requires more time than direct touch interaction for 1D cutting, 2D drilling and 3D extraction. While it seems that the position of the referent (horizontal or vertical) slightly influenced the 3D extraction time, we found that there is almost no difference for 1D cutting and

2D drilling. Finally, for any of the three tasks, the vertical or horizontal position of the physical referent does not impact the results in terms of error rate.

Furthermore, when comparing these results involving standing participants with those obtained in previous studies (with seated participants), it appears that the user's position (standing or seated) does not impact the task completion time or error rate. Overall, whether the user is standing or seated, using a horizontal or vertical physical referent, we found that mid-air interaction requires more time than direct touch on the physical referent to perform selection tasks on a STC.

## 9 DISCUSSION AND FUTURE WORK

In this section, we first summarize the core outcomes from our studies. We then describe the key principles of the embedded interaction presented in this paper. Then, we analyze how our techniques could be deployed, the limitations of our work and anticipate future work.

### 9.1 Synthesis and analysis of our results

Our work establishes that direct touch interaction on a physical referent is more efficient and appreciated than the traditional mid-air interaction for the three STC data analysis tasks considered: 1D cutting, 2D drilling and 3D extraction. Indeed, our preliminary study revealed that touching directly the edges or faces of the physical referent with the finger, provides the user with a direct mapping between the interaction space and the resulting manipulation and seems, therefore, more appropriate to them. This result is underlined by participants stating that *"I feel like I'm touching the faces of the space-time cube"*. Furthermore, touching the referent provides a support for interaction, thus reducing fatigue and offering a physical delimitation of the interaction space. Regarding the physical features, the preliminary study also established that participants prefer using the physical edges of the referent for 1D cutting and its faces for 2D drilling and 3D extraction (with a mode-switching mechanism to disambiguate the use of the face when required). As expressed above, participants interacted on the physical referent as they would have interacted with the space-time cube itself: touching a 1D edge for 1D cutting, 2D face for 2D drilling, etc.

In addition, the results of our first study show that direct touch on the physical referent allows for faster 1D and 2D selection, reduces fatigue and is preferred over mid-air interaction. We believe that these results are explained by the physical referent properties: 1) The hand is laid on the referent, thus reducing fatigue; 2) The participant can place its finger on a specific spot of an edge or face, physically delimited by the referent shape, thus allowing for faster initial position of the cursor (i.e. start the selection directly at  $x = 20$  instead of starting at 50 then moving the cursor to 20); 3) Finally, the physical borders of the faces and edges allow the user to perceive the minimum and maximum bounds of the interaction space. However, for 3D extraction, mid-air interaction performs better than direct touch. Indeed, 3D extraction with touch interaction requires the user to touch two faces (front for XY and top for Z) in sequence for each selection step (two faces for tens and two faces for units). We believe that this two-step process is the origin of the increase in completion time with our 3D extraction techniques. We thus iterated on the design of our direct touch interaction technique: we introduced bi-manual interaction to allow for simultaneous interaction on two faces (front and left) of the physical referent.

Our second study results confirm that combining the two sequential face selection processes into a bimanual simultaneous two-face selection increases the touch interaction performance and preference, in line with previous works on bimanual touch-based interaction [90]. Furthermore, this technique still benefits from the advantages of

touch interaction over mid-air interaction (discoverability, knowing in advance the interaction space bounds, reducing fatigue).

Furthermore, our third study results show that the position of the physical referent (horizontal or vertical) does not impact the performance or error rate: the physical referent was placed at the same height regarding the user (height of the elbow), whether it was horizontally or vertically fixed. As such, there was no difference in terms of interaction with the physical referent, which explains the similarity between the results in both conditions. Finally, when comparing these results with our previous studies, it appears that the user's position (standing or seated) has no impact on the technique completion time or error rate. We conclude that using a stable referent (i.e. a referent fixed on a surface) allows for efficient tactile interaction, independently of the way the referent is fixed and of the user's posture during interaction.

## **9.2 From embedded representations to embedded interactions**

The context of our work is related to how to use a physical referent to support multidimensional data selection on a data visualization (such as a space-time cube) displayed in AR. This context is challenging due to the large number of possible mappings between the input modes, physical referent features, STC dimensions and analysis tasks. The originality of our approach lies in combining and extending the advantages of 1) direct touch interaction, 2) embedded data representations and 3) augmented reality visualizations.

- 1) Direct touch interaction is beneficial for learning, manipulation and eyes-free interaction. Our work explores if these benefits are still valid in a novel situation where the physical referent cannot be manipulated because it is anchored (either fixed or wired, for example) in its environment.
- 2) Embedded data representations facilitate interaction with data by tightly coupling the physical object (the referent) to which the data is related with its representation. In this work, we adopt this approach by displaying the data close to the physical referent using AR. In addition, our approach further emphasizes the inner link between the physical referent and its embedded data representation by using the physical referent as a support for input interaction with the data representation (a STC in our work).
- 3) AR visualizations have been largely explored because of their almost unlimited display space, which offers many possibilities for representing large information spaces. In our work, we address one of the main issues of AR environments, i.e. input interaction, by proposing an ecological alternative to mid-air gestures (tiring and challenging to discover) or dedicated external input controllers (cumbersome when on the go). Our solution opportunistically exploits an already integrated object into the user's activity environment: the physical referent.

Our approach is symmetrical to embedded visualization: instead of enriching the physical world with visual projections, our work enhances a digital representation by using the physical world object to which the data is related.

Consequently, we believe that embedded data visualizations, combined with a physical referent to support the interaction, is an exciting interaction paradigm. As we are unaware of other works specifically describing and characterizing such approach in the literature, we propose to name it "phygital" interaction as a reminder of its combination of PHYsical interaction with a diGITAL data representation.



### 9.3 Deploying “phygital interaction”

In our exploration, we used an OptiTrack tracking system to detect the hand position and the touch interaction on the physical referent. This implementation provides high tracking precision and speed, ensuring the internal validity of our experimental results, but is less well suited for real-world deployment as it requires users to wear markers on their fingers. To move towards a more ecological deployment, recent work [54], such as Capstones [23], Capricate [93], or Sensurfaces [80], integrate touch detection directly within their materials. Another option we have already started prototyping is to use the tracking capabilities of HoloLens 2 to detect the user’s finger position and the Vuforia Model Tracker [113] to detect the position and orientation in the space of the physical referent. While this solution allows interacting with any physical model, it would require the user to look at the model while interacting because the volume in which HoloLens2 tracks the hand corresponds approximately to its field of view. An external head-mounted camera pointing downwards could solve this issue. We will further work on such real-world deployment in the future.

### 9.4 Limitations and future work

Our work is only a first step toward using a physical referent to support the interaction with a digital data representation. In this section, we describe the limitations of our studies and how they could be addressed through future works.

Our studies highlighted the effectiveness of touch interaction with a physical referent in terms of speed and user preference. We plan to conduct a longitudinal study to better understand the advantages of such interaction in terms of memorization, which are already well-established for tangible and touch interaction [107]. It could also be relevant to associate physical loci (i.e. pre-defined physical location) directly on the physical referent. This would facilitate access to more than the seven tasks considered in this work (1D cutting on X, Y, and Z axes; 2D drilling on XY, XZ, and YZ planes and 3D extraction). This approach could be relevant to interact with other tasks or commands such as filtering, rotating/translating the STC, etc. For instance, a touch on a corner of the physical referent could activate a command selection mode in which all available commands are projected on one face of the physical referent. The user could then navigate through these commands, select one by touching its physical loci on the referent, and activate an interaction mode to execute this specific command on the physical referent. Another solution, inspired by [39], would be to use different fingers or multiple fingers on the STC depending on the desired task. For example, one finger would correspond to selections on the STC. Two fingers would allow rotating/translating the STC. Three fingers would provide a way to activate commands such as changing the axes of the data visualization or filtering on a specific range.

Another interesting parameter to explore will be the size of our physical referent. Indeed, performing touch interaction on an edge or a face of the physical referent may be less accurate on a smaller surface (<3cm) or would take too much time on a very large surface (>20cm). Moreover, faces with very different widths and heights (for example, a 3x20cm surface) might confuse the user as the precision required for gestures in one direction is higher than in the other direction.

Furthermore, extending our work to other shapes of the physical referent would be relevant. We chose to use a cubic-shaped physical referent as it is a standard shape for everyday objects providing data (internet box, electric meter, etc.) and as a first exploration step of embedded data analysis. However, Conic or cylinder-based shapes are also fairly common in everyday environments (e.g. a cup of coffee, a lamp base). Our paper presents a first step that will later need to be adapted to these specific shapes.

Finally, while we describe two distinct usage scenarios taking advantage of multidimensional data selection in Augmented Reality with a physical referent as input (see section 3.1), we envisioned that they could be extended to other pertinent contexts of use (smart homes, industrial contexts, crisis management). It would be interesting to explore how our interaction techniques would enhance in-situ data access through longitudinal studies in one of these contexts.

## 10 CONCLUSION

In this paper, we explored using a physical referent to facilitate the input interaction with a STC displayed in AR. Our work focuses on STC selection tasks (1D cutting, 2D drilling, 3D extraction). In a preliminary study, we explored the possible mappings to achieve these selections using the referent's physical features (corners, edges, faces). Based on these results, we designed 3 interaction techniques that we compared to a baseline involving mid-air gestures in a first study. Our results show that direct touch interaction on the physical referent requires less time, is less tiring and more appreciated than mid-air interaction for 1D cutting and 2D drilling tasks. Mid-air interaction takes up to 88% more time than touch interaction on the physical referent for 1D Cutting and 35% more time for 2D Drilling. Following a design iteration on our touch interaction techniques for the 3D extraction task, we proposed two new bimanual interaction techniques that we compared to mid-air interaction in a second study. Our results show that using bimanual touch on the physical referent requires less time, is less tiring and more appreciated than the mid-air interaction of the HoloLens: mid-air interaction takes up to 38% more time than a bimanual-touch technique on the physical referent. Finally, we explored the interaction with the physical referent when the user is standing in front of it either horizontally or vertically fixed on a surface. Our results show that even in such settings, mid-air interaction still requires more time than direct touch interaction. Combining all of our studies' results, we found that direct touch interaction on the physical referent always requires less time, is preferred, has a higher usability score and induces less fatigue than mid-air interaction. These results remain valid for any of our selection tasks (1D cutting, 2D drilling and 3D extraction), whether the user is seated or standing and, when the user is standing, whether the physical referent is horizontally or vertically fixed on a surface.

## ACKNOWLEDGMENTS

This work was lead as part of the Vilagil-AmUrb, INTERPLAY and JCJC PERFIN projects. Vilagil-AmUrb is supported by the French Government in the framework of the major investment plan (Territoires d'Innovation), by Toulouse Métropole and the neOCampus Scientific Interest Group of the University of Toulouse III Paul Sabatier. INTERPLAY is supported by the french research agency ANR (grant ANR-21-CE33-0022). JCJC PERFIN is supported by the french research agency ANR (grant ANR-18-CE33-0009).

## REFERENCES

- [1] Khalid W. Al-Ani, Ahmed Subhi Abdalkafor, and Aiman Majid Nassar. 2019. Smart City Applications: A Survey. In Proceedings of the 9th International Conference on Information Systems and Technologies (icist 2019). Association for Computing Machinery, New York, NY, USA, Article 36, 1–4. DOI:<https://doi.org/10.1145/3361570.3361616>
- [2] AR Urban Design, HitLabNZ ([https://www.youtube.com/watch?v=5\\_zE7bahRD0](https://www.youtube.com/watch?v=5_zE7bahRD0) last visited 20/02/2023)
- [3] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 185, 15 pages. DOI: <https://doi.org/10.1145/3173574.3173759>
- [4] Takumi Azai, Shuhei Ogawa, Mai Otsuki, Fumihisa Shibata, and Asako Kimura. 2017. Selection and Manipulation Methods for a Menu Widget on the Human Forearm. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17). Association for Computing Machinery, New York, NY, USA, 357–360.

DOI:<https://doi.org/10.1145/3027063.3052959>

- [5] B. Bach, R. Sicat, J. Beyer, M. Cordeil and H. Pfister, "The Hologram in My Hand: How Effective is Interactive Exploration of 3D Visualizations in Immersive Tangible Augmented Reality?," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 457-467, Jan. 2018, doi: 10.1109/TVCG.2017.2745941.
- [6] Benjamin Bach, Pierre Dragicevic, Daniel Archambault, Christophe Hurter, and Sheelagh Carpendale. 2016. A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes. In *Computer Graphics Forum*. Wiley Online Library.
- [7] Benjamin Bach, Pierre Dragicevic, Daniel Archambault, Christophe Hurter, Sheelagh Carpendale. A Review of Temporal Data Visualizations Based on Space-Time Cube Operations. Eurographics Conference on Visualization, Jun 2014, Swansea, Wales, United Kingdom. fhal-01006140
- [8] Benjamin Bach, Emmanuel Pietriga, and Jean-Daniel Fekete. 2014. Visualizing Dynamic Networks with Matrix Cubes. In *Proceedings of SIGCHI Conference on Human Factors and Computing Systems (CHI '14)*. ACM, New York, NY, USA. to appear.
- [9] Sumathi Balakrishnan, Hemalata Vasudavan, and Raja Kumar Murugesan. 2018. Smart Home Technologies: A Preliminary Review. In *Proceedings of the 6th International Conference on Information Technology: IoT and Smart City (ICIT 2018)*. Association for Computing Machinery, New York, NY, USA, 120–127. DOI:<https://doi.org/10.1145/3301551.3301575>
- [10] A. Batch et al., "There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 536-546, Jan. 2020, doi: 10.1109/TVCG.2019.2934803.
- [11] Patrick Baudisch, Nathaniel Good, and Paul Stewart. 2001. Focus Plus Context Screens: Combining Display Technology with Visualization Techniques. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST '01)*, 31–40. DOI: <https://doi.org/10.1145/502348.502354>
- [12] L. Besançon, P. Issartel, M. Ammi and T. Isenberg, "Hybrid Tactile/Tangible Interaction for 3D Data Exploration," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 881-890, Jan. 2017, doi: 10.1109/TVCG.2016.2599217.
- [13] Lonni Besançon, Paul Issartel, Mehdi Ammi, and Tobias Isenberg. 2017. Mouse, Tactile, and Tangible Input for 3D Manipulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4727–4740. DOI:<https://doi.org/10.1145/3025453.3025863>
- [14] Oriël Bergig, Eyal Soreq, Nate Hagbi, Kirill Pevzner, Nati Levi, Shoham Blau, Yulia Smelansky, Jihad El-Sana, "Out of the Cube: Augmented Rubik's Cube", *International Journal of Computer Games Technology*, vol. 2011, Article ID 570210, 7 pages, 2011. <https://doi.org/10.1155/2011/570210>
- [15] Billinghamurst, M., Kato, H., & Poupyrev, I. (2008). *Tangible augmented reality*. SIGGRAPH 2008.
- [16] Borg, G. Borg's Perceived Exertion and Pain Scales. *Human Kinetics* (1998), viii 104pp.
- [17] E. Brasier, O. Chapuis, N. Ferey, J. Vezien and C. Appert, "ARpads: Mid-air Indirect Input for Augmented Reality," 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2020, pp. 332-343, doi: 10.1109/ISMAR50242.2020.00060.
- [18] R. Budhiraja, G. A. Lee and M. Billinghamurst, "Using a HMD with a HMD for mobile AR interaction," 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2013, pp. 1-6, doi: 10.1109/ISMAR.2013.6671837.
- [19] Wolfgang Büschel, Annett Mitschick, Thomas Meyer, and Raimund Dachsel. 2019. Investigating Smartphone-based Pan and Zoom in 3D Data Spaces in Augmented Reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 2, 1–13. DOI:<https://doi.org/10.1145/3338286.3340113>
- [20] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: Multi-"Touch" Interaction Around Small Devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*, 201–204. DOI: <https://doi.org/10.1145/1449715.1449746>
- [21] M. Sheelagh T. Carpendale, D. J. Cowperthwaite, M. Tigges, A. J. Fall, F. David Fracchia, "Tardis: a visual exploration environment for landscape dynamics," *Proc. SPIE 3643, Visual Data Exploration and Analysis VI*, (25 March 1999); <https://doi.org/10.1117/12.342825>
- [22] Arpan Chakraborty, Ryan Gross, Shea McIntee, Kyung Wha Hong, Jae Yeol Lee, and Robert St. Amant. 2014. CAPTIVE: a cube with augmented physical tools. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 1315–1320. DOI:<https://doi.org/10.1145/2559206.2581340>
- [23] Liwei Chan, Stefanie Müller, Anne Roudaut, and Patrick Baudisch. 2012. CapStones and ZebraWidgets: sensing stacks of building blocks, dials and sliders on capacitive touch screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2189–2192. DOI:<https://doi.org/10.1145/2207676.2208371>
- [24] Quijano-Chavez C., Nedel L., Freitas C.M.D.S. (2021) An Immersive Approach Based on Two Levels of Interaction for Exploring Multiple Coordinated 3D Views. In: Ardito C. et al. (eds) *Human-Computer Interaction – INTERACT 2021*. INTERACT 2021. Lecture Notes in Computer Science, vol 12934. Springer, Cham. [https://doi.org/10.1007/978-3-030-85613-7\\_33](https://doi.org/10.1007/978-3-030-85613-7_33)
- [25] Cityscope: <https://www.media.mit.edu/projects/cityscope/overview/> (Last visited 20/02/2023)
- [26] M. Cordeil, B. Bach, A. Cunningham, B. Montoya, R. T. Smith, B. H. Thomas, and T. Dwyer. Embodied Axes: Tangible, Actuated Interaction for 3D Augmented Reality Data Spaces. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376613>
- [27] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. 2017. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. Association for Computing Machinery, New York, NY, USA, 71–83.

DOI:<https://doi.org/10.1145/3126594.3126613>

- [28] Maxime Cordeil, Benjamin Bach, Yongchao Li, Elliott Wilson, and Tim Dwyer. 2017. Design space for spatio-data coordination: Tangible interaction devices for immersive information visualisation. In 2017 IEEE Pacific Visualization Symposium (PacificVis). IEEE, 46–50.
- [29] David J. Cowperthwaite, M. Sheelagh T. Carpendale, and F. David Fracchia. 1996. Visual access for 3D data. In Conference Companion on Human Factors in Computing Systems (CHI '96). Association for Computing Machinery, New York, NY, USA, 175–176. DOI:<https://doi.org/10.1145/257089.257242>
- [30] Sebastian Doeweling, Tarik Tahiri, Philipp Sowinski, Benedikt Schmidt, and Mohammadreza Khalilbeigi. 2013. Support for collaborative situation analysis and planning in crisis management teams using interactive tabletops. In Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces (ITS '13). Association for Computing Machinery, New York, NY, USA, 273–282. <https://doi.org/10.1145/2512349.2512823>
- [31] Niklas Elmqvist, Pierre Dragicevic, and Jean-Daniel Fekete. 2008. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on Visualization and Computer Graphics* 14, 6 (2008), 1539–1148.
- [32] Neven A.M. Elsayed, Bruce H. Thomas, Ross T. Smith, Kim Marriott, and Julia Piantadosi. 2015. Using augmented reality to support situated analytics. In 2015 IEEE Virtual Reality Conference, VR 2015 - Proceedings. 175–176. <https://doi.org/10.1109/VR.2015.7223352>
- [33] Oumaima El Rhazal and Mazri Tomader. 2019. Study of smart city data: categories and quality challenges. In Proceedings of the 4th International Conference on Smart City Applications (SCA '19). Association for Computing Machinery, New York, NY, USA, Article 4, 1–7. DOI:<https://doi.org/10.1145/3368756.3368965>
- [34] D. Englmeier, J. Dörner, A. Butz and T. Höllerer, "A Tangible Spherical Proxy for Object Manipulation in Augmented Reality," 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2020, pp. 221-229, doi: 10.1109/VR46266.2020.00041.
- [35] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafer, Matt Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 459, 1–17. DOI:<https://doi.org/10.1145/3411764.3446866>
- [36] Barrett Ens and Pourang Irani. 2017. Spatial analytic interfaces: Spatial user interfaces for in situ visual analytics. *IEEE Computer Graphics and Applications* 37, 2 (2017), 66–79. <https://doi.org/10.1109/MCG.2016.38>
- [37] Tiago Gonçalves, Ana Paula Afonso, and Bruno Martins. 2014. Visualizing Human Trajectories: Comparing Space-Time Cubes and Static Maps. In Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI (BCS-HCI '14). BCS, Swindon, GBR, 207–212. <https://doi.org/10.14236/ewic/hci2014.24>
- [38] Ens B, Goodwin S, Prouzeau A, Anderson F, Wang FY, Gratzl S, Lucarelli Z, Moyle B, Smiley J, Dwyer T. Uplift: A Tangible and Immersive Tabletop System for Casual Collaborative Visual Analytics. *IEEE Trans Vis Comput Graph.* 2021 Feb;27(2):1193-1203. doi: 10.1109/TVCG.2020.3030334. Epub 2021 Jan 28. PMID: 33074810.
- [39] Aakar Gupta, Muhammed Anwar, and Ravin Balakrishnan. 2016. Porous Interfaces for Small Screen Multitasking using Finger Identification. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16). Association for Computing Machinery, New York, NY, USA, 145–156. DOI:<https://doi.org/10.1145/2984511.2984557>
- [40] Sidney Fels, Eric Lee, and Kenji Mase. 2000. Techniques for interactive video cubism (poster session). In Proceedings of the eighth ACM international conference on Multimedia (MULTIMEDIA '00). Association for Computing Machinery, New York, NY, USA, 368–370. DOI:<https://doi.org/10.1145/354384.354535>
- [41] George W. Fitzmaurice. 1993. Situated information spaces and spatially aware palmtop computers. *Commun. ACM* 36, 7 (July 1993), 39-49. DOI: <http://doi.org/10.1145/159544.159566>
- [42] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. 2013. inFORM: dynamic physical affordances and constraints through shape and object actuation.. In *Uist*, Vol. 13. 417–426.
- [43] Bernd Fröhlich and John Plate. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '00). Association for Computing Machinery, New York, NY, USA, 526–531. DOI:<https://doi.org/10.1145/332040.332491>
- [44] Gronier, G. & Baudet, A. (2021). Psychometric evaluation of the F-SUS: Creation and validation of the French version of the System Usability Scale. *International Journal of Human-Computer Interaction.* <https://doi.org/10.1080/10447318.2021.1898828>
- [45] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: Multi Fidelity Interaction with Displays On and Around the Body. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), 3933-3942. DOI: <https://doi.org/10.1145/2702123.2702331>
- [46] Aakar Gupta, Bo Rui Lin, Sinyi Ji, Arjav Patel, and Daniel Vogel. 2020. Replicate and Reuse: Tangible Interaction Design for Digitally-Augmented Physical Media Objects. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–12. DOI:<https://doi.org/10.1145/3313831.3376139>
- [47] Devamardeep Hayatpur, Haijun Xia, and Daniel Wigdor. 2020. DataHop: Spatial Data Exploration in Virtual Reality. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20). Association for Computing Machinery, New York, NY, USA, 818–828. DOI:<https://doi.org/10.1145/3379337.3415878>
- [48] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of midair interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, Association

for Computing Machinery, New York, NY, USA, 1063–1072

- [49] Mikkel R. Jakobsen, Yvonne Jansen, Sebastian Boring, and Kasper Hornbæk. 2015. Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction. In 15th IFIP TC 13 International Conference on Human-Computer Interaction --- INTERACT 2015 - Volume 9298. Springer-Verlag, Berlin, Heidelberg, 455–473. DOI:[https://doi.org/10.1007/978-3-319-22698-9\\_31](https://doi.org/10.1007/978-3-319-22698-9_31)
- [50] Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling Cumulative Arm Fatigue in Mid-Air Interaction Based on Perceived Exertion and Kinetics of Arm Motion. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 3328–3339. <https://doi.org/10.1145/3025453.3025523>
- [51] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 3227–3236. DOI:<https://doi.org/10.1145/2702123.2702180>
- [52] Y. Jansen, P. Dragicevic, and J.-D. Fekete. Tangible remote controllers for wall-size displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, p. 2865–2874. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2207676.2208691
- [53] Thorsten Karrer, Moritz Wittenhagen, Leonhard Lichtschlag, Florian Heller, and Jan Borchers. 2011. Pinstripe: eyes-free continuous input on interactive clothing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, NY, USA, 1313-1322. DOI: <https://doi.org/10.1145/1978942.1979137>
- [54] Kunihiro Kato, Kaori Ikematsu, and Yoshihiro Kawahara. 2020. CAPath: 3D-Printed Interfaces with Conductive Points in Grid Layout to Extend Capacitive Touch Inputs. Proc. ACM Hum.-Comput. Interact. 4, ISS, Article 193 (November 2020), 17 pages. DOI:<https://doi.org/10.1145/3427321>
- [55] Nurit Kirshenbaum, James Hutchison, Ryan Theriot, Dylan Kobayashi, and Jason Leigh. 2020. Data in Context: Engaging Audiences with 3D Physical Geo-Visualization. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–9. DOI:<https://doi.org/10.1145/3334480.3382968>
- [56] M. J. Kraak. 2003. The Space-Time Cube revisited from a Geovisualization Perspective. Proceedings of the International Cartographic Conference (2003), 1988–1996.
- [57] Krzywinski, M., Altman, N.: points of significance: error bars. Nat. Methods 10(10), 921– 922 (2013)
- [58] Eun Kwon, Gerard J. Kim, and Sangyoon Lee. 2009. Effects of sizes and shapes of props in tangible augmented reality. In Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '09). IEEE Computer Society, USA, 201–202. DOI:<https://doi.org/10.1109/ISMAR.2009.5336463>
- [59] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Paper 81, 1–14. DOI:<https://doi.org/10.1145/3173574.3173655>
- [60] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. 2021. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 468, 1–17. DOI:<https://doi.org/10.1145/3411764.3445593>
- [61] Khanh-Duy Le, Tanh Quang Tran, Karol Chlasta, Krzysztof Krejtz, Morten Fjeld, and Andreas Kunz. 2021. VXSlate: Exploring Combination of Head Movements and Mobile Touch for Large Virtual Display Interaction. In Designing Interactive Systems Conference 2021 (DIS '21). Association for Computing Machinery, New York, NY, USA, 283–297. DOI:<https://doi.org/10.1145/3461778.3462076>
- [62] D. Li, D. Weng, Y. Li and J. Xie, "TouchableAR: A New Experience of Augmented Reality," 2013 International Conference on Virtual Reality and Visualization, Xi'an, 2013, pp. 37-42, doi: 10.1109/ICVRV.2013.14.
- [63] Zhen Li, Joannes Chan, Joshua Walton, Hrvoje Benko, Daniel Wigdor, and Michael Glueck. 2021. Armstrong: An Empirical Examination of Pointing at Non-Dominant Arm-Anchored UIs in Virtual Reality. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 123, 1–14. DOI:<https://doi.org/10.1145/3411764.3445064>
- [64] Linky: <https://www.enedis.fr/le-compteur-linky-un-outil-pour-la-transition-ecologique> (last visited 20/02/2023)
- [65] Lars Lischke, Paweł W. Woźniak, Sven Mayer, Andreas Preikschat, and Morten Fjeld. 2017. Using Variable Movement Resistance Sliders for Remote Discrete Input. In Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17). ACM, New York, NY, USA, 116–125. DOI: <http://dx.doi.org/10.1145/3132272.3134135>
- [66] I. Scott MacKenzie. 2013. Human-Computer Interaction: An Empirical Research Perspective. Morgan Kaufmann is an imprint of Elsevier, Amsterdam.
- [67] MATLAB. 2010. version 7.10.0 (R2010a). The MathWorks Inc., Natick, Massachusetts
- [68] Vande Moere A., Patel S. (2009) The Physical Visualization of Information: Designing Data Sculptures in an Educational Context. In: Huang M., Nguyen Q., Zhang K. (eds) Visual Information Communication. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4419-0312-9\\_1](https://doi.org/10.1007/978-1-4419-0312-9_1)
- [69] A. V. Moere, "Beyond the Tyranny of the Pixel: Exploring the Physicality of Information Visualization," 2008 12th International Conference Information Visualisation, 2008, pp. 469-474, doi: 10.1109/IV.2008.84.
- [70] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. 2019. TrackCap: Enabling

- Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Paper 585, 11 pages. DOI: <https://doi.org/10.1145/3290605.3300815>
- [71] Rafael Morales González, Caroline Appert, Gilles Bailly, and Emmanuel Pietriga. 2016. TouchTokens: Guiding Touch Patterns with Passive Tokens. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16). Association for Computing Machinery, New York, NY, USA, 4189–4202. DOI:<https://doi.org/10.1145/2858036.2858041>
- [72] Mama N. Mouchili, Shadi Aljawarneh, and Wette Tchouati. 2018. Smart city data analysis. In Proceedings of the First International Conference on Data Science, E-learning and Information Systems (DATA '18). Association for Computing Machinery, New York, NY, USA, Article 33, 1–6. DOI:<https://doi.org/10.1145/3279996.3280029>
- [73] MRTK: <https://github.com/microsoft/MixedRealityToolkit-Unity> (last visited 20/02/2023)
- [74] Annov K Mutasim, Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Short Papers). Association for Computing Machinery, New York, NY, USA, Article 15, 1–7. DOI:<https://doi.org/10.1145/3448018.3457998>
- [75] Mathieu Nancel, Olivier Chapuis, Emmanuel Pietriga, Xing-Dong Yang, Pourang P. Irani, and Michel Beaudouin-Lafon. 2013. High-precision pointing on large wall displays using small handheld devices. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 831–840. DOI:<https://doi.org/10.1145/2470654.2470773>
- [76] NeoCampus: <https://www.irit.fr/neocampus/fr/> (last visited 20/02/2023)
- [77] Erwan Normand and Michael J. McGuffin. 2018. Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display). In 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 123–133. DOI:<https://doi.org/10.1109/ISMAR.2018.00043>
- [78] Claire O'Malley, Danae Stanton Fraser. Literature Review in Learning with Tangible Technologies. 2004. (hal-00190328)
- [79] Optitrack: <https://optitrack.com/> (last visited 20/02/2023)
- [80] Brice Parilusyian, Marc Teyssier, Valentin Martinez-Missir, Clément Duhart, and Marcos Serrano. 2022. Sensurfaces: A Novel Approach for Embedded Touch Sensing on Everyday Surfaces. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 6, 2, Article 67 (July 2022), 19 pages. <https://doi.org/10.1145/3534616>
- [81] Daniela Petrelli, Alessandro Soranzo, Luigina Ciolfi, and John Reidy. 2016. Exploring the Aesthetics of Tangible Interaction: Experiments on the Perception of Hybrid Objects. In Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '16), 100–108. <https://doi.org/10.1145/2839462.2839478>
- [82] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. In 25th ACM Symposium on Virtual Reality Software and Technology (VRST '19). Association for Computing Machinery, New York, NY, USA, Article 35, 1–11. DOI:<https://doi.org/10.1145/3359996.3364264>
- [83] Henning Pohl, Klemen Lilija, Jess McIntosh, and Kasper Hornbæk. 2021. Poros: Configurable Proxies for Distant Interactions in VR. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 532, 1–12. DOI:<https://doi.org/10.1145/3411764.3445685>
- [84] Arnaud Prouzeau, Yuchen Wang, Barrett Ens, Wesley Willett, and Tim Dwyer. 2020. Corsican Twin: Authoring in Situ Augmented Reality Visualisations in Virtual Reality. In ACM International Conference Proceeding Series. <https://doi.org/10.1145/3399715.3399743>
- [85] VandenBos, G.R. (ed.): Publication Manual of the American Psychological Association, 6th edn. American Psychological Association, Washington, DC (2009). <http://www.apastyle.org/manual/>
- [86] Anke Verena Reinschluessel, Thomas Muender, Tanja Döring, Verena Nicole Uslar, Thomas Lück, Dirk Weyhe, Andrea Schenk, and Rainer Malaka. 2021. A Study on the Size of Tangible Organ-shaped Controllers for Exploring Medical Data in VR. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 261, 1–7. DOI:<https://doi.org/10.1145/3411763.3451594>
- [87] P. Reipschlager, T. Flemisch and R. Dachselt, "Personal Augmented Reality for Information Visualization on Large Interactive Displays," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 1182-1192, Feb. 2021, doi: 10.1109/TVCG.2020.3030460.
- [88] Houssein Saidi, Emmanuel Dubois, and Marcos Serrano. 2021. HoloBar: Rapid Command Execution for Head-Worn AR Exploiting Around the Field-of-View Interaction. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 745, 1–17. DOI:<https://doi.org/10.1145/3411764.3445255>
- [89] Houssein Saidi, Marcos Serrano, Pourang Irani, Christophe Hurter, and Emmanuel Dubois. 2018. On-Body tangible interaction: using the body to support tangible manipulations for immersive visualization. In Proceedings of the 30th Conference on l'Interaction Homme-Machine (IHM '18). Association for Computing Machinery, New York, NY, USA, 1–11. DOI:<https://doi.org/10.1145/3286689.3286692>
- [90] Houssein Saidi, Marcos Serrano, and Emmanuel Dubois. 2016. Investigating the effects of splitting detailed views in Overview+Detail interfaces. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16). Association for Computing Machinery, New York, NY, USA, 180–184. <https://doi.org/10.1145/2935334.2935341>
- [91] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 483–492. DOI:<https://doi.org/10.1145/2207676.2207743>

- [92] Kim Sauvé, Dominic Potts, Jason Alexander, and Steven Houben. 2020. A Change of Perspective: How User Orientation Influences the Perception of Physicalizations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI:<https://doi.org/10.1145/3313831.3376312>
- [93] Martin Schmitz, Mohammadreza Khalilbeigi, Matthias Balwierz, Roman Lissermann, Max Mühlhäuser, and Jürgen Steimle. 2015. Capricate: A Fabrication Pipeline to Design and 3D Print Capacitive Touch Sensors for Interactive Objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. Association for Computing Machinery, New York, NY, USA, 253–258. DOI:<https://doi.org/10.1145/2807442.2807503>
- [94] S. A. Seah, B. W. Drinkwater, T. Carter, R. Malkin and S. Subramanian, "Correspondence: Dexterous ultrasonic levitation of millimeter-sized objects in air," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, no. 7, pp. 1233–1236, July 2014, doi: 10.1109/TUFFC.2014.3022.
- [95] Marcos Serrano, Dale Hildebrandt, Sriram Subramanian, and Pourang Irani. 2014. Identifying suitable projection parameters and display configurations for mobile true-3D displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services (MobileHCI '14)*. ACM, New York, NY, USA, 135–143. DOI: <https://doi.org/10.1145/2628363.2628375>
- [96] Marcos Serrano, Barrett M. Ens, and Pourang P. Irani. 2014. Exploring the use of hand-to-face input for interacting with head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3181–3190. DOI:<https://doi.org/10.1145/2556288.2556984>
- [97] Aubrey Simonson. 2021. Pockets: User-Assigned Menus Based on Physical Buttons for Virtual Environments. In *ACM SIGGRAPH 2021 Posters (SIGGRAPH '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 1–2. DOI:<https://doi.org/10.1145/3450618.3469174>
- [98] Orit Shaer and Eva Hornecker (2010), "Tangible User Interfaces: Past, Present, and Future Directions", *Foundations and Trends® in Human-Computer Interaction*: Vol. 3: No. 1–2, pp 4–137. <http://dx.doi.org/10.1561/1100000026>
- [99] Smart home controller : [https://www.bosch-smarthome.com/fr/fr/produits/appareils/controleur-bosch-smarthome/?gclid=CjwKCAiAvaGRBhBIeIwAiY-yM0cz4MhZvZzXEpSznVsYjVCjT\\_jsKLHDlwwjlu0\\_uF5Y1uogMYjnRoC8Z8QAvD\\_BwE](https://www.bosch-smarthome.com/fr/fr/produits/appareils/controleur-bosch-smarthome/?gclid=CjwKCAiAvaGRBhBIeIwAiY-yM0cz4MhZvZzXEpSznVsYjVCjT_jsKLHDlwwjlu0_uF5Y1uogMYjnRoC8Z8QAvD_BwE) (last visited 20/02/2023)
- [100] Smart industry: <https://www.plm.automation.siemens.com/global/en/products/manufacturing-planning/plant-simulation-throughput-optimization.html> (last visited 20/02/2023)
- [101] Jim Smiley, Benjamin Lee, Siddhant Tandon, Maxime Cordeil, Lonni Besançon, Jarrod Knibbe, Bernhard Jenny, and Tim Dwyer. 2021. The MADE-Axis: A Modular Actuated Device to Embody the Axis of a Data Dimension. *Proc. ACM Hum.-Comput. Interact.* 5, ISS, Article 501 (November 2021), 23 pages. DOI:<https://doi.org/10.1145/3488546>
- [102] Christian Steins, Sean Gustafson, Christian Holz, and Patrick Baudisch. 2013. Imaginary devices: gesture-based interaction mimicking traditional input devices. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13)*. Association for Computing Machinery, New York, NY, USA, 123–126. DOI:<https://doi.org/10.1145/2493190.2493208>
- [103] Sophie Stellmach and Raimund Dachselt. 2013. Still looking: investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 285–294. DOI:<https://doi.org/10.1145/2470654.2470695>
- [104] Sophie Stellmach and Raimund Dachselt. 2012. Look & touch: gaze-supported target acquisition. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2981–2990. DOI:<https://doi.org/10.1145/2207676.2208709>
- [105] Junwei Sun, Wolfgang Stuerzlinger, and Bernhard E. Riecke. 2018. Comparing input methods and cursors for 3D positioning with head-mounted displays. In *Proceedings of the 15th ACM Symposium on Applied Perception (SAP '18)*. Association for Computing Machinery, New York, NY, USA, Article 8, 1–8. DOI:<https://doi.org/10.1145/3225153.3225167>
- [106] Faisal Taher, Yvonne Jansen, Jonathan Woodruff, John Hardy, Kasper Hornbaek, and Jason Alexander. 2017. Investigating the Use of a Dynamic Physical Bar Chart for Data Exploration and Presentation. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (January 2017), 451–460. DOI:<https://doi.org/10.1109/TVCG.2016.2598498>
- [107] F. Taher, J. Hardy, A. Karnik, C. Weichel, Y. Jansen, K. Hornbaek, and J. Alexander. Exploring Interactions with Physically Dynamic Bar Charts. In *CHI 2015 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3237–3246. Seoul, South Korea, Apr. 2015. doi: 10.1145/2702123.2702604
- [108] Tang, S. K., Sekikawa, Y., Leithinger, D., Follmer, S., and Ishii, H. Tangible cityscape, 2013. [tinyurl.com/tcityscape](http://tinyurl.com/tcityscape).
- [109] Philip Tuddenham, David Kirk, and Shahram Izadi. 2010. Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2223–2232. DOI:<https://doi.org/10.1145/1753326.1753662>
- [110] VILAGIL: <https://metropole.toulouse.fr/actualites/vilagil-favoriser-de-nouvelles-formes-de-mobilites> (last visited 20/02/2023)
- [111] Simon Voelker, Andrii Matviienko, Johannes Schöning, and Jan Borchers. 2015. Combining Direct and Indirect Touch Input for Interactive Workspaces using Gaze Input. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction (SUI '15)*. Association for Computing Machinery, New York, NY, USA, 79–88. DOI:<https://doi.org/10.1145/2788940.2788949>

- [112] Daniel Vogel and Géry Casiez. 2011. Conté: multimodal input inspired by an artist's crayon. In Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11). Association for Computing Machinery, New York, NY, USA, 357–366. <https://doi.org/10.1145/2047196.2047242>
- [113] Vuforia Model Tracker: <https://library.vuforia.com/features/objects/model-targets.html> (last visited 20/02/2023)
- [114] Yun Wang, Adrien Segal, Roberta Klatzky, Daniel F. Keefe, Petra Isenberg, Jörn Hurtienne, Eva Hornecker, Tim Dwyer, Stephen Barrass, and Theresa-Marie Rhyne. 2019. An Emotional Response to the Value of Visualization. *IEEE Computer Graphics and Applications* 39 (2019), 8–17.
- [115] Colin Ware. 2008. *Information Visualization - Perception for Design*. Vol. 41. Morgan Kaufmann. CR – Copyright © 2008 Leonardo pages. arXiv:arXiv:1011.1669v3 <http://www.jstor.org/stable/20206579>
- [116] Frederik Wiehr, Florian Daiber, Felix Kosmalla, and Antonio Krüger. 2017. ARTopos: augmented reality terrain map visualization for collaborative route planning. In Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17). Association for Computing Machinery, New York, NY, USA, 1047–1050. DOI:<https://doi.org/10.1145/3123024.3124446>
- [117] Daniel Wigdor, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. 2007. Lucid Touch: A See-through Mobile Device. In Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07), 269–278. DOI: <https://doi.org/10.1145/1294211.1294259>
- [118] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2017. Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (January 2017), 461–470. DOI:<https://doi.org/10.1109/TVCG.2016.2598608>
- [119] Willett, W., Jenny, B., Isenberg, T., et al.: Lightweight relief shearing for enhanced terrain perception on interactive maps. In: Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI 2015) (2015). <https://doi.org/10.1145/2702123.2702172>
- [120] R. Xiao, J. Schwarz, N. Throm, A. D. Wilson and H. Benko, "MRTouch: Adding Touch Input to Head-Mounted Mixed Reality," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 4, pp. 1653-1660, April 2018, doi: 10.1109/TVCG.2018.2794222.
- [121] Yiwei Zhao, Lawrence H. Kim, Ye Wang, Mathieu Le Goc, and Sean Follmer. 2017. Robotic Assembly of Haptic Proxy Objects for Tangible Interaction and Virtual Reality. In Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17). Association for Computing Machinery, New York, NY, USA, 82–91. DOI:<https://doi.org/10.1145/3132272.3134143>
- [122] Jack Zhao and Andrew Vande Moere. 2008. Embodiment in data sculpture: a model of the physical visualization of information. In Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts (DIMEA '08). Association for Computing Machinery, New York, NY, USA, 343–350. DOI:<https://doi.org/10.1145/1413634.1413696>
- [123] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14. DOI:<https://doi.org/10.1145/3313831.3376233>
- [124] Müller-Tomfelde C. (2007) Dwell-Based Pointing in Applications of Human Computer Interaction. In: Baranauskas C., Palanque P., Abascal J., Barbosa S.D.J. (eds) *Human-Computer Interaction – INTERACT 2007*. INTERACT 2007. Lecture Notes in Computer Science, vol 4662. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-74796-3\\_56](https://doi.org/10.1007/978-3-540-74796-3_56)

## APPENDICES

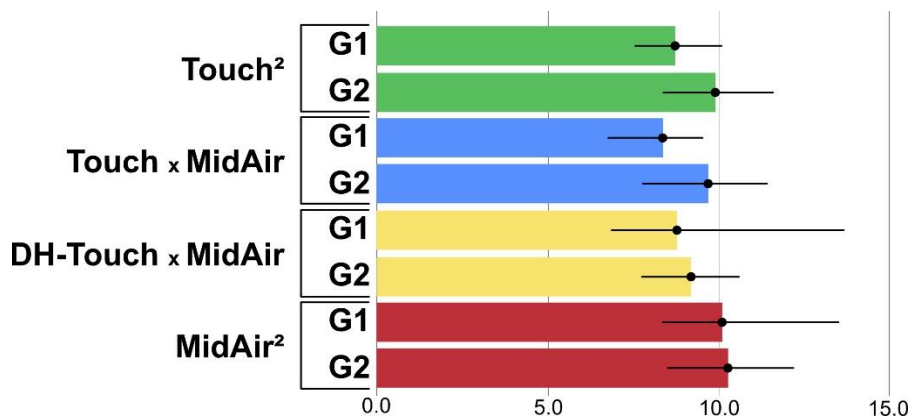


Figure 18. Average completion time (all tasks) for each technique and each user group in our Study 1 (CI 95%). G1 represents users who participated in our preliminary study. G2 represents users who did not.



Technique	CI_Min	Mean time	CI_Max
G1_DH-Touch x MidAir	6.88	8.81	13.71
G2_DH-Touch x MidAir	7.76	9.21	10.63
G1_Touch <sup>2</sup>	7.57	8.75	10.10
G2_Touch <sup>2</sup>	8.39	9.93	11.62
G1_MidAir <sup>2</sup>	8.36	10.12	13.56
G2_MidAir <sup>2</sup>	8.53	10.30	12.22
G1_Touch x MidAir	6.78	8.39	9.56
G2_Touch x MidAir	7.79	9.73	11.44

Table 1. Detailed completion time (all tasks) for each technique and user group shown in Figure 14.

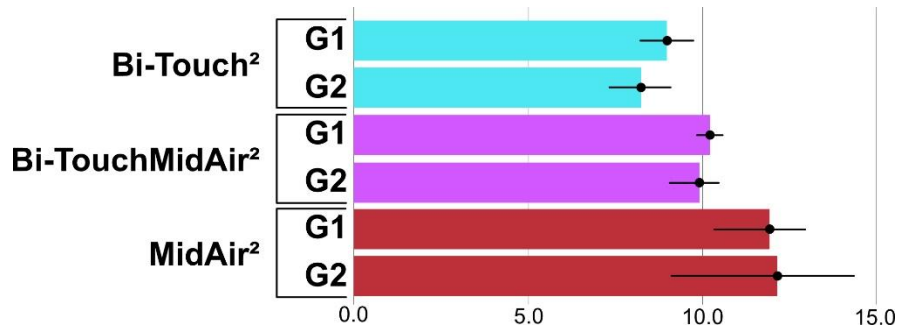


Figure 19. Average completion time for each technique and each user group in our Study 2 (CI 95%). G1 represents users who participated in our study 1. G2 represents users who did not.

Technique	CI_Min	Mean time	CI_Max
G1_Bi-Touch <sup>2</sup>	8.25	9.02	9.79
G2_Bi-Touch <sup>2</sup>	7.36	8.28	9.12
G1_MidAir <sup>2</sup>	10.37	11.97	12.99
G2_MidAir <sup>2</sup>	9.15	12.19	14.39
G1_Bi-TouchMidAir <sup>2</sup>	9.87	10.26	10.63
G2_Bi-TouchMidAir <sup>2</sup>	9.07	9.96	10.52

Table 2. Detailed completion time (all tasks) for each technique and user group shown in Figure 15.

Technique	CI_Min	Mean time	CI_Max
Touch <sup>2</sup>	3.85	4.19	4.49
Touch x MidAir	4.43	5.17	5.98
DH-Tactile x MidAir	4.56	5.10	6.44
MidAir <sup>2</sup>	5.52	6.17	7.18

Table 3. Detailed completion time without menu time for 1D cutting in Figure 9.

Technique	CI_Min	Mean time	CI_Max
MidAir <sup>2</sup> / Touch <sup>2</sup>	1.37	1.44	1.61
MidAir <sup>2</sup> / Touch x MidAir	1.09	1.23	1.35
MidAir <sup>2</sup> / DH-Touch x MidAir	1.11	1.26	1.48

Table 4. Detailed completion time (ratio) without menu time for 1D cutting in Figure 9.

Technique	CI_Min	Mean time	CI_Max
Touch <sup>2</sup>	6.97	8.07	9.54
Touch x MidAir	6.89	8.10	9.27
DH-Tactile x MidAir	7.53	9.34	14.62
MidAir <sup>2</sup>	7.73	9.17	11.14

Table 5. Detailed completion time without menu time for 2D drilling in Figure 10.

Technique	CI_Min	Mean time	CI_Max
MidAir <sup>2</sup> / Touch <sup>2</sup>	1.02	1.15	1.37
MidAir <sup>2</sup> / Touch x MidAir	1.02	1.13	1.27
MidAir <sup>2</sup> / DH-Touch x MidAir	0.94	1.05	1.18

Table 6. Detailed completion time (ratio) without menu time for 2D Drilling in Figure 10.

Technique	CI_Min	Mean time	CI_Max
Bi-Touch <sup>2</sup>	7.31	7.98	8.56
Bi-TouchMidAir <sup>2</sup>	8.32	9.69	10.27
MidAir <sup>2</sup>	8.94	10.7	12.22

Table 7. Detailed completion time without menu time for 3D extraction in Figure 13.

Technique	CI_Min	Mean time	CI_Max
MidAir <sup>2</sup> / Bi-Touch <sup>2</sup>	1.19	1.31	1.44
MidAir <sup>2</sup> / Bi-TouchMidAir <sup>2</sup>	0.98	1.08	1.19

Table 8. Detailed completion time (ratio) without menu time for 3D extraction in Figure 13.

Technique	CI_Min	Mean time	CI_Max
Horizontal-MidAir <sup>2</sup>	5.37	6.16	7.30
Vertical-MidAir <sup>2</sup>	4.85	5.76	7.63

Table 9. Detailed completion time without menu time for 1D cutting with Horizontal and Vertical physical referents in Figure 15.

Technique	CI_Min	Mean time	CI_Max
Horizontal: MidAir <sup>2</sup> / Touch <sup>2</sup>	1.29	1.40	1.56
Vertical: MidAir <sup>2</sup> / Touch <sup>2</sup>	1.10	1.28	1.50

Table 10. Detailed completion time (ratio) without menu time for 1D cutting with Horizontal and Vertical physical referents in Figure 15.

Technique	CI_Min	Mean time	CI_Max
Horizontal-MidAir <sup>2</sup>	7.09	8.47	7.80
Vertical-MidAir <sup>2</sup>	7.42	9.30	13.99

Table 11. Detailed completion time without menu time for 2D drilling with Horizontal and Vertical physical referents in Figure 16.

Technique	CI_Min	Mean time	CI_Max
Horizontal: MidAir <sup>2</sup> / Touch <sup>2</sup>	1.06	1.16	1.30
Vertical: MidAir <sup>2</sup> / Touch <sup>2</sup>	1.05	1.21	1.65

Table 12. Detailed completion time (ratio) without menu time for 2D drilling with Horizontal and Vertical physical referents in Figure 16.

Technique	CI_Min	Mean time	CI_Max
Horizontal-MidAir <sup>2</sup>	10.24	11.71	13.41
Vertical-MidAir <sup>2</sup>	9.54	11.33	13.12

Table 13. Detailed completion time without menu time for 3D extraction with Horizontal and Vertical physical referents in Figure 17.

Technique	CI_Min	Mean time	CI_Max
Horizontal: MidAir <sup>2</sup> / Bi-Touch <sup>2</sup>	1.09	1.19	1.29
Vertical: MidAir <sup>2</sup> / Bi-Touch <sup>2</sup>	1.23	1.35	1.48

Table 14. Detailed completion time (ratio) without menu time for 3D extraction with Horizontal and Vertical physical referents in Figure 17.