



## A Tangible Volume for Portable 3D Interaction

Paul Issartel, Lonni Besançon, Tobias Isenberg, Mehdi Ammi

### ► To cite this version:

Paul Issartel, Lonni Besançon, Tobias Isenberg, Mehdi Ammi. A Tangible Volume for Portable 3D Interaction. 2016. hal-01284985

**HAL Id: hal-01284985**

**<https://hal.science/hal-01284985>**

Preprint submitted on 8 Mar 2016

**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.

# A Tangible Volume for Portable 3D Interaction

Paul Issartel\*  
LIMSI-CNRS  
Univ. Paris-Sud

Lonni Besançon†  
INRIA Saclay  
Univ. Paris-Sud

Tobias Isenberg†  
INRIA Saclay

Mehdi Ammi\*  
LIMSI-CNRS  
Univ. Paris-Sud

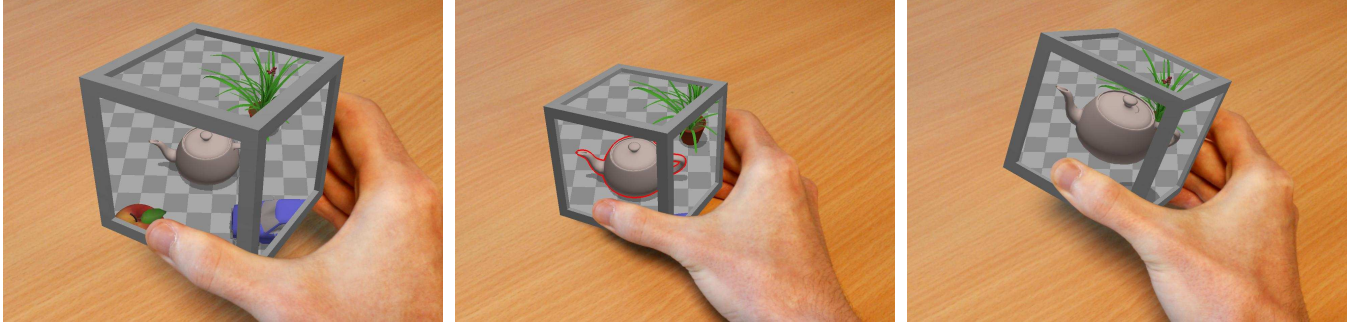


Figure 1: Our concept of a *tangible volume* consists of a fully portable and self-contained device, entirely covered with screens. A virtual scene can be seen “through” the volume of the device (left image). This volume can be directly positioned within the virtual scene (middle image), and used to grasp and manipulate virtual objects (right image). A cubic-shaped device is shown here for illustration, but other volume shapes can be used.

## ABSTRACT

We present a new approach to achieve tangible object manipulation with a single, fully portable and self-contained device. Our solution is based on the concept of a “tangible volume”. We turn a tangible object into a handheld fish-tank display. The tangible volume represents a volume of space that can be freely manipulated within a virtual scene. This volume can be positioned onto virtual objects to directly grasp them, and to manipulate them in 3D space. We investigate this concept through two user studies. The first study evaluates the intuitiveness of using a tangible volume for grasping and manipulating virtual objects. The second study evaluates the effects of the limited field of view on spatial awareness. Finally, we present a generalization of this concept to other forms of interaction through the surface of the volume.

**Keywords:** Tangible User Interface, 3D Manipulation, Fish-Tank VR

**Index Terms:** H.5.1 [Information Interfaces And Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

## 1 INTRODUCTION

Interaction with 3D data, and in particular object manipulation [3] (selecting, translating and rotating 3D virtual objects) is of major importance in many fields, such as scientific visualization, prototyping and gaming. Until recently, these tasks were often carried out by advanced users on fixed workstations. However, with the development of mobile computing, people are now expecting to perform these tasks anywhere, with minimal set-up and learning. Thus, there is a need for an interface that is not only *natural and efficient* for 3D manipulation, but also truly *portable and self-contained*.

Tangible User Interfaces (TUIs) represent a promising approach for 3D manipulation. They consist of physical objects, or tangible objects, that serve as real-world representations of digital data. In many TUIs, tangible objects are used as physical “handles” for virtual objects. These interfaces take advantage of the user’s skills

in interacting with physical objects [14], making them an attractive solution to support manipulation tasks in a natural and efficient way.

However, interacting with virtual objects also requires visual feedback. Tangible UIs for 3D manipulation often rely on an external and fixed monitor for visual output [10]. Hence, such interfaces cannot be considered portable. Others use mobile devices as a portable display surface [16]. Although these interfaces are portable, they still consist of multiple separate pieces, which always have to be handled and carried together. They are not self-contained. Some TUIs take a further step and use the mobile display as a physical handle [9]. The device itself replaces the external tangible objects, thereby reducing the entire interface to a single portable and self-contained object. This solution, however, does not eliminate the distance between the display and the virtual objects. Since the mobile display now serves as the tangible handle, this separation creates problems during manipulation, such as a shifted center of rotation or manipulated objects leaving the field of view.

In this paper, we present a different approach. Rather than turning a mobile device into a tangible handle, we turn a tangible object into a display. We cover the surface of a tangible object with multiple screens, and use fish-tank rendering to display part of the virtual scene “through” the object. The object then becomes a tangible representation of a *volume* of virtual space. We introduce a “grasping” metaphor that consists in positioning this volume directly onto a virtual object, and pressing the fingers to pick up the object with the tangible volume. Thus, the separation between the tangible handle and the manipulated virtual object disappears. By solving the issues caused by this separation, our solution makes it possible to preserve the advantages of tangible manipulation in a fully portable and self-contained device. Our contributions in this work are therefore:

- a new concept of a portable and self-contained device that allows users to directly interact with a volume of virtual space,
- a grasping metaphor that demonstrates an application of this concept to 3D manipulation,
- a partial prototype and two experimental studies to investigate several usability aspects of this mode of interaction, and
- a generalization of our concept to other forms of interaction through the surface of the volume.

\*e-mail: {paul.issartel,mehdi.ammi}@limsi.fr

†e-mail: {lonni.besancon,tobias.isenberg}@inria.fr

## 2 BACKGROUND

### 2.1 Tangible User Interfaces for 3D manipulation

The main idea behind Tangible User Interfaces (TUIs) is to give physical form to digital data [14, 15]. This is accomplished through the use of real-world objects—called *tangible objects*—that represent the digital information. One particular type of TUI are Graspable User Interfaces [6], in which tangible objects serve as physical “handles” for virtual objects. Each handle can be attached to a virtual object. Users can then manipulate a virtual object by directly manipulating the corresponding tangible object. By taking advantage of the user’s preexisting skills in manipulating real-world objects [6, 14], this interaction mode provides a natural and immediately efficient approach to 3D manipulation.

A number of TUIs have been designed around this concept. Hinckley et al. [10] used passive tangible objects (“props”) tracked in mid-air to position and orient 3D medical data. The result was displayed on a separate computer screen. Similar interfaces were later proposed by Qi and Martens [24] and Jackson et al. [17] with wireless tracking. The Personal Space Station [22] is a situated virtual reality system that consists of a projector and a reflective screen. Users interact with the virtual 3D objects by manipulating tangible objects behind the screen. A limitation of these interfaces, however, is that they all require an external monitor or projector. This makes them essentially fixed installations. The lack of portability creates many constraints. Users need to go to a dedicated place in order to use the system. It cannot be carried between offices or brought home. Nowadays, with the rise of mobile computing, such constraints appear even more limiting. There is a need for a truly portable interface that would offer the same advantages as fixed TUIs. Thus, our first requirement is that the entire interface should be *portable*.

These constraints can be addressed by making the display portable. Since mobile devices—such as smartphones and tactile tablets—have become widespread, they represent a readily-available solution for use as a portable display surface. Some researchers have combined mobile device with tangible objects [1, 16, 18], resulting in portable TUIs. However, these interfaces also consist of multiple independent pieces: the mobile device and the tangible objects, all of which need to be carried together at all times. Whenever one of the tangible objects is missing, or when the mobile device is separated from the tangible objects, the interface becomes less functional or even unusable. The portability advantage is reduced due to the inconvenience of having to carry and keep the multiple pieces together. We can thus identify another requirement: the interface should not only be portable but should also consist of a single, *self-contained* device.

### 2.2 Mobile device as tangible handle

In addition to its role as a portable display surface, the mobile device itself can be seen as a tangible object. Several researchers have proposed to use the mobile device as tangible handle [9, 20, 21] to manipulate virtual objects in the space behind the device (generally in augmented reality). This eliminates the need for external tangible props, resulting in an interface that is fully portable and self-contained in a single object: the mobile device.

Although this solution fully meets the portability requirements, it is not, however, without drawbacks. Unlike previous approaches, there is now a separation between the virtual objects (behind the device) and the tangible handle (the device itself). This separation leads to problems during manipulation. In the techniques proposed by Henrysson et al. [9] and Marzo et al. [20], the manipulated object is fixed relative to the device. Since the center of rotation is now located on the device, it becomes difficult to rotate a virtual object without also translating it (Figure 2(a)). In the HOMER-S technique [21], translations and rotations are separately applied to the manipulated object, which makes it easier to rotate the object about itself. On the other hand, the object is no longer fixed relative to the mobile device and can thus leave the field of view during

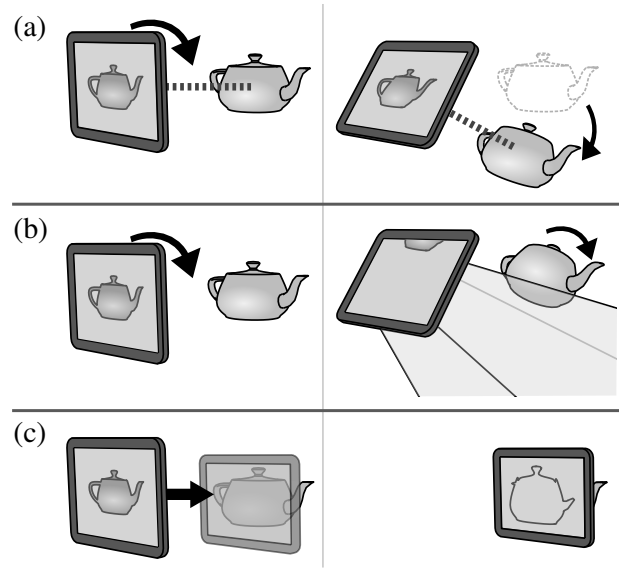


Figure 2: Manipulation issues that arise when a typical mobile device is used as a tangible handle: (a) the virtual object does not rotate about its own center; (b) the virtual object leaves the field of view; (c) when trying to reduce the distance to avoid the above issues, the virtual object is clipped.

rotations (Figure 2(b)). The authors suggest to alter the control-display ratio to avoid this limitation, but such an indirect mapping would become less natural than direct tangible manipulation.

One way to completely eliminate the separation, and its associated problems, would be to position the mobile device *onto* the virtual object before starting the manipulation. However, a typical rendering process would clip half of the virtual object in this situation—as well as everything else in front of the screen—making it difficult to manipulate the object properly (Figure 2(c)). Additionally, most current mobile devices come in the form of flat one-sided displays. Even if clipping was not an issue, this form factor would still result in loss of visual feedback during manipulation. Therefore, we can identify two additional requirements: the portable and self-contained device should provide visual feedback on its entire surface, and should be able to display virtual objects without clipping.

### 2.3 Fish-Tank VR and geometric displays

When rendering a 3D scene, the display is generally considered as the viewpoint on the virtual world. This causes clipping whenever a virtual object crosses the screen plane. Fish-Tank Virtual Reality (FTVR) provides a solution to address the clipping problem. FTVR turns the display into a “window” seen from the viewpoint of the user [28]. This allows virtual objects to appear behind, onto or in front of the display surface [7]. There are two main ways to achieve this effect: stereo rendering, head-coupled perspective, or a combination of both [28]. According to Ware et al. [28], head coupling is a much stronger clue than stereo rendering, making it the preferred solution to implement this technique.

As explained before, a flat one-sided display is not the best shape for use as a tangible handle. A number of works have proposed volumetric objects equipped with multiple displays, capable of providing visual feedback on their entire surface. Some of these objects take advantage of FTVR to give the illusion of a 3D space inside the device. They are called *geometric displays* [26]. One of the first examples is the MEDIA cube [13]: multiple LCD displays were arranged in a box shape, and combined with head coupling to give the illusion that a virtual scene was inside the box. The CoCube [4]

is a tangible cube that produced the same illusion when seen through a head-mounted display (HMD). Unlike the previous example, this tangible cube was freely manipulable by the user. The Cubee [27] is a cubic device that achieved the same goal without a HMD, by using integrated displays and a head tracker. The pCubee [26] is an evolution of Cubee that was made smaller and more portable. A similar device is gCubik [19], which uses autostereoscopic rendering instead of head tracking. Most of the above devices have a cuboid shape, as this is the easiest way to arrange conventional rectangular displays, but other shapes are possible—such as arbitrary polyhedra [8] or even a sphere [5].

Geometric displays appear to meet all our requirements: they can be made small, portable and self-contained, they provide visual feedback on all their surface, and with FTVR they can display virtual objects without clipping. However, previous work on this subject was mainly focused on the feasibility of such displays. Although significant technological contributions have been made, the potential of such devices for 3D interaction remains comparatively little explored. For example, the pCubee device [26] is one of the few works in which interaction techniques are mentioned—yet those are limited to visualizing 3D objects and tumbling them inside the display, while object selection requires an external input device. In addition, there seems to be a lack of examples of positioning the geometric display itself within a larger virtual scene, which is an essential part of our concept. Even though the pCubee can be used to navigate in the virtual scene, this is accomplished through an indirect velocity-based mapping. In contrast, our concept is based on a full 1-1 mapping between the real world and the virtual world, which has greater potential for direct 3D interaction. In this work, we offer a new perspective on geometric displays, by considering such a device as a portable and self-contained TUI in its own right, capable of both displaying a virtual scene and serving as a tangible handle within the scene itself.

### 3 CONCEPT: A TANGIBLE VOLUME

Based on the above considerations, we introduce the concept of a “tangible volume”. A tangible volume is a single physical object, sufficiently small and lightweight to be held in the hand. The surface of this object is entirely covered in screens on which the virtual scene is displayed. The perspective of each screen is adjusted to the user’s head position. As a result, part of the virtual scene appears “through” the object. The object is also tracked relative to the real world. Users can reach other parts of the virtual scene by moving the object in real space (Figure 3). At any point, the physical boundaries of the object “enclose” a corresponding part of the virtual scene. Therefore, this object is a *tangible* representation of a *volume* of virtual space, that can be held in the hand and directly positioned into the virtual scene.

We use this tangible volume as an interaction device for 3D object manipulation. First, the tangible volume is positioned *onto* a virtual object in the scene. From there, the virtual object is attached to the volume. The virtual object then follows the tangible volume in 3D space, as if it was directly held in the hand (Figure 1). Thus, there is no separation between the virtual object and the tangible handle.

Our concept integrates all input and output into a single handheld object, used both to visualize the virtual scene and to manipulate virtual objects. Unlike the alternative approach of using a mobile device as a tangible handle, our interface also eliminates the separation from the manipulated virtual objects. Our concept thus constitutes a fully portable and self-contained interface for 3D object manipulation which avoids the problems described before.

#### 3.1 Object selection by grasping

Having an interface made of an unique tangible object has clear benefits for portability, but also raises the question of how to interact with more than one virtual object. Such an interface must

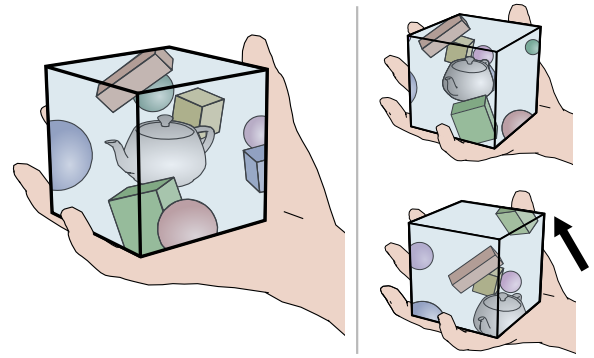


Figure 3: Illustration of our concept of a “tangible volume”. Part of a larger virtual scene can be seen through the tangible object held by the user. On the top right, the tangible volume is observed from a different angle. On the bottom right, the tangible volume has been translated in space and now encloses a different part of the virtual scene.

provide a way to attach and detach virtual objects from the single tangible handle.

In TUIs made of multiple tangible objects, each tangible object can be linked to a different virtual object. This allows the user to interact easily with multiple virtual objects, simply by manipulating the corresponding tangible objects as desired. This is called “space multiplexing” [6]. In our case, there is only one tangible object available to manipulate an arbitrary number of virtual objects. Therefore, the user has to *select* which virtual object should be linked to the tangible object at any given time. This is called “time multiplexing” [6]. Space multiplexing is considered more desirable than time multiplexing, due to the lack of an object selection step which improves efficiency and lowers cognitive load [6].

However, let us consider more closely what selection means in the specific case of tangible objects. Even though some interfaces allow the user to interact with multiple tangible objects concurrently, this is ultimately limited by the human capabilities. Tangible manipulation is typically performed with the hand(s). This means that all manipulation is accomplished through at most two effectors: the hands themselves. Any tangible object first has to be grasped with the hand in order to interact with the corresponding virtual object. Therefore, there is still an implicit selection in space-multiplexed TUIs that occurs when an object is grasped with the hand.

From that perspective, one of the main advantages of space multiplexed TUIs is not actually the lack of a selection step, but rather the fact this selection is implicit and does not require thinking. In other words, reaching for an object and grasping it with the hand constitutes a *natural* way to select it.

#### Grasping metaphor

The tangible volume provides a unique opportunity to reproduce this form of selection with virtual objects: by considering the volume as an extension of the hand. As shown above, the tangible volume can be moved in space and positioned around a virtual object. Since the volume is held in the hand, and is surrounding a virtual object, the hand is also surrounding the virtual object.

We thus designed a selection technique that consists in *pressing the fingers* on the tangible volume to “grasp” the virtual object located inside (Figure 4). Because the tangible volume is already held in the hand, the grasping technique is only triggered when finger pressure exceeds a given threshold. This threshold allows to differentiate a conscious grasping action from normal manipulation of the tangible volume itself. Similarly, releasing the virtual object is done by releasing finger pressure below this threshold, while keeping hold of the tangible volume.

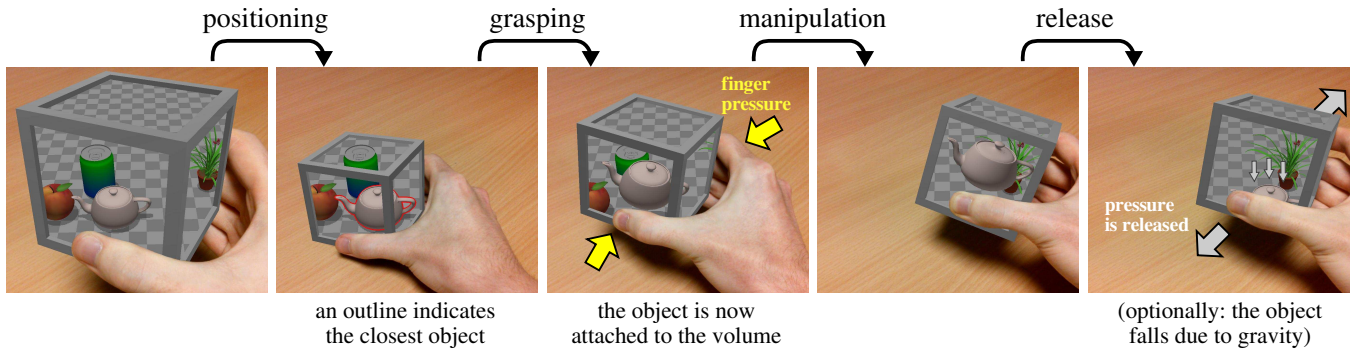


Figure 4: Illustration of the different steps of object manipulation in our interface concept. First, the tangible volume is positioned onto a virtual object. To disambiguate between nearby objects, an outline indicates which object is both inside the volume and closest to its center. This object can be grasped by pressing the fingers on the volume. The object is then attached to the volume, and can be directly moved alongside the volume in 3D space. The object is detached when finger pressure is released. If virtual gravity is enabled, it then falls to the ground.

With our grasping metaphor, picking up a virtual object becomes very similar to picking up a real-world object: placing the hand around an object, and pressing the fingers to grasp it.

### 3.2 Disambiguating between virtual objects

A challenge that arises from our grasping metaphor is that fingers cannot penetrate the volume. When virtual objects are close to each other, more than one may be located inside the volume bounds. Since fingers cannot be directly used to grasp the desired object, there must be a way to indicate which object will be selected among those located in the volume. One possible solution is to display an outline around the object that is closest to the center of the volume (Figure 4).

### 3.3 Bimanual manipulation

Space-multiplexed TUIs allow the user to manipulate different virtual objects in each hand. Since our interface is made of a single object, this form of bimanual manipulation is not directly supported. One solution could be to use two tangible volumes—one for each hand. Since our grasping metaphor makes the tangible volume an extension of the hand, and all manipulation is accomplished through the hands, this solution would approximate space-multiplexed bimanual interaction. However, the interface would then again consist of multiple pieces, which we specifically wanted to avoid.

Still, as a single object, our interface supports another form of bimanual interaction: manipulating a single virtual object with both hands. This can be accomplished by holding the tangible volume with two hands, and pressing the respective fingers on opposite sides of the volume.

### 3.4 Simulated physics

In his dissertation on Graspable User Interfaces, Fitzmaurice [6], citing Norman [23], argues that the natural laws of physics that affect tangible objects help the user during manipulation. Indeed, in graspable UIs, releasing an object from the hand makes it drop to the floor due to gravity. This familiar behavior can help understand and predict its motion.

However, in our interface, the tangible volume always remains in the hand. Releasing a virtual object simply detaches the virtual object from the volume. Without further intervention, the virtual object would remain there—floating in space. We can obtain a more realistic behavior by adding simulated physics to the virtual scene. With physics, the released virtual object falls to the virtual ground when released (Figure 4). Simulated physics also prevents virtual objects from moving through each other. This reinforces the illusion that they are solid, and thus can be grasped and manipulated directly.

It should be noted that virtual physics is not always desirable. For example, a complex manipulation task may have to be decomposed

in several steps in mid-air (clutching). With virtual gravity, releasing the virtual object between each step would cause it to drop from its intermediate positions. Virtual physics enhances realism, but is not necessarily appropriate for every application.

## 4 IMPLEMENTATION

Having presented our concept as we envision it, we discuss here its technical feasibility given the current state of technology. We also describe a partial prototype, used in the two studies presented in the next sections.

### 4.1 Technical feasibility

Our concept has three main technological requirements. The first requirement is to cover the surface of a volumetric object with screens, as well as pressure sensors. Since the device is meant to be portable and self-contained, it should also contain enough processing capabilities to render a 3D scene on multiple screens at interactive frame-rates, with enough battery power to operate for a sufficient duration. In view of the current advancement of geometric displays, as demonstrated by previous work, and the rapid pace of technological progress in the mobile device industry, this first aspect appears very close to be achievable.

The second requirement is to track the user’s head position relative to the device, which is important to produce the fish-tank effect. In our concept, tracking requirements are made more challenging by the fact the device should be self-contained. Previous work on portable geometric displays (e.g. [26]) relied on magnetic trackers, one attached to the user’s head and the other to the object itself, along with an external tracking base. Assuming the base can be embedded in the object or in the head-worn device, and the head-worn device is comfortable and unobtrusive enough (e.g. a pair of glasses), this solution represents a good compromise given the current state of technology. For a truly self-contained device, small cameras could be embedded on the surface of the object in order to optically track the user’s head. Although current optical head tracking would not be reliable enough in such a configuration, this solution remains theoretically achievable in the future.

The third requirement is to track the device relative to the real world. This is needed to produce the illusion that the virtual scene is fixed relative to the real world, which allows users to perform all the manipulation in real space. In a truly self-contained interface, there cannot be any external tracker to provide the real-world reference frame. Thus, the only solution is to use inside-out tracking from the device itself. Recently, the Tango project<sup>1</sup> demonstrated how a device can track its own motion relative to the environment with

<sup>1</sup><https://www.google.com/atap/project-tango/>

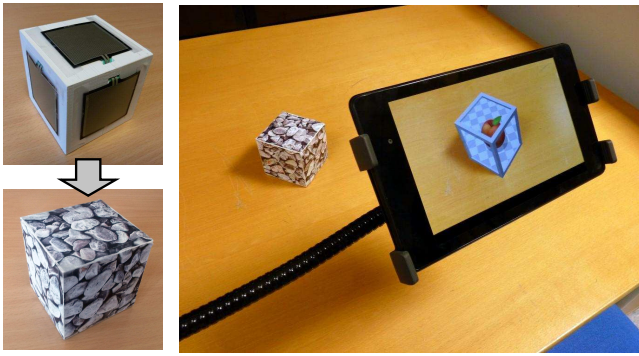


Figure 5: A partially simulated implementation of our concept, using augmented reality. A physical cube serves as the tangible volume. We covered its faces with six pressure sensors, then with AR markers. A tablet, raised to eye level, turns the markers into virtual screens. Users manipulate the cube by looking through the tablet, as if they were directly holding the real device.

an integrated infrared camera. Several of these cameras could be embedded on the surface of the device in order to provide continuous tracking of its real-world motion.

## 4.2 Partial prototype

In order to study several usability aspects of our concept, we designed a preliminary prototype. As said before, our work primarily focuses on the interaction capabilities afforded by a tangible volume, rather than on the hardware side. Therefore, we used augmented reality (AR) to simulate some hardware aspects of our concept.

One may argue that existing geometric displays could have served as a basis to conduct user studies, without requiring AR simulation. However, existing implementations described in previous work were all lacking in key aspects of our concept. Some of these devices are fixed installations, not meant to be moved by the user [13]. Others are somewhat movable, but are still tethered to a larger workstation [19, 26]. The tangible volume as intended in our concept is supposed to be fully portable and self-contained, and a tether wire could hinder manipulation [11]. Finally, some of these devices are only partially covered in screens [26, 27]. Even though this is sufficient to demonstrate the technology, for complex object manipulation (especially rotations) it is important that all sides of the volume provide visual feedback. Simulating an interaction device to conduct user studies has been done in previous work [2]. This approach has the advantage of providing flawless rendering and tracking, which would be difficult to achieve in a research prototype but essential for the validity of user studies.

We chose a cubic shape for the tangible volume in our prototype. Even though other shapes could have been used, we chose a cube because it was easier to build while also being easy to manipulate by users. This cube was covered with AR markers. We used a tactile tablet as an augmented reality window. When observed through the rear camera of the tablet, the faces of the cube were replaced with “virtual screens” that displayed part of the virtual scene (Figure 5). A frame was added around each virtual screen, to account for the fact that real screens would not be completely borderless. Using AR provided implicit viewpoint tracking: since the cube faces were tracked by the camera, reversing this transformation produced an equivalent result. The cube was tracked relative to the real world by placing an additional AR marker in the environment. We employed the Vuforia framework<sup>2</sup> to track these objects. The tablet was attached to a raised stand, so that users could simply manipulate the object

<sup>2</sup><http://www.vuforia.com/>

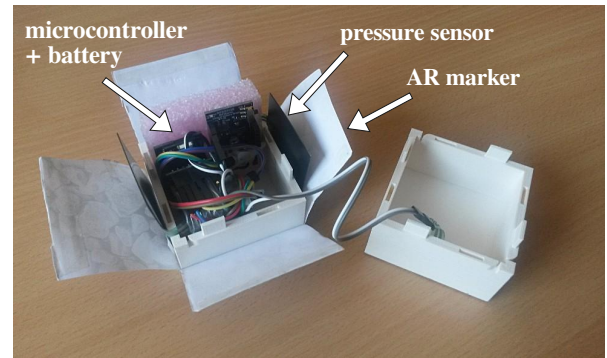


Figure 6: View of the electronic components inside the cube. The embedded microcontroller retrieves values from the pressure sensors on the cube surface, and sends them to the rendering software on the tablet through a wireless connection.

behind the tablet, as if they were directly looking at a cube equipped with screens.

Our grasping technique was implemented in actual hardware, by attaching six flat pressure sensors (Interlink® FSR 406) to the faces of the tangible cube (Figure 5). The sensors were located under the AR markers, and thus were invisible to the users. They were driven by a microcontroller<sup>3</sup> embedded in the cube (Figure 6). It was powered by a rechargeable battery, with a charging port hidden in one corner of the cube. The microcontroller continuously transmitted the pressure values to the rendering software on the tablet, through a wireless Bluetooth connection, at a frequency of 10 Hz. Finally, we implemented physical simulation for the grasped objects with the Bullet<sup>4</sup> physics engine.

In the rest of this paper, we use this prototype to investigate two important questions about our concept. First, we evaluate the intuitiveness of object manipulation with our grasping metaphor. In a second study, we evaluate the effects of the small field of view, which is limited to the surface of the tangible volume, on spatial awareness.

## 5 USER STUDY: OBJECT MANIPULATION

When interacting with the real world, grasping an object by pressing fingers on it, and moving it while maintaining finger pressure is a fairly natural procedure. However, it is unknown whether the exact same procedure remains natural when interacting with virtual objects through a tangible volume. We thus conducted an experiment to evaluate the *intuitiveness* of object selection and manipulation in our interface.

More specifically, we wanted to see if users can understand *by themselves* with no prior explanations, how to grasp and move virtual objects with our interface. Of course, if some users do not succeed without help, this measure alone would not be sufficient to understand why. Therefore, we also wanted to determine at which step of manipulation those users would need guidance on their initial encounter with the interface.

### 5.1 Participants and apparatus

This study was conducted with unpaid 36 participants (from 20 to 52 years old, mean=29.5, sd=9.4). None of them had any prior knowledge of our interface concept or our grasping metaphor.

The apparatus was our AR-based prototype described in the previous sections. We slightly raised the pressure threshold to ensure that the grasping technique would only be triggered when truly intended by the participants.

<sup>3</sup>RFduino RFD22102

<sup>4</sup><http://www.bulletphysics.org/>

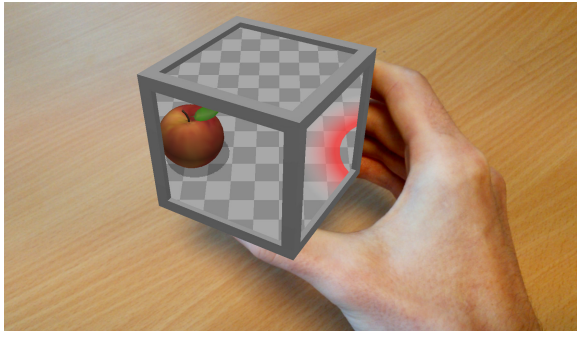


Figure 7: Screenshot of our first experiment, as seen by participants through the tablet. Note the lack of finger occlusion, due to the AR simulation. The task was to pick up the virtual apple and move it to the red circle on the right.

## 5.2 Procedure

Before starting the experiment, we first told participants that it would consist in “manipulating the cube”. We then explained why there was also a tablet in front of them: to “simulate screens that should have been on the faces of the cube, but were not there due to technical limitations”. We followed by demonstrating how the cube turned into a cubic display when placed behind the tablet, and how a virtual scene could be seen through it. If our prototype had actual screens and head tracking, participants would likely have noticed all of this immediately. We did this short demonstration to ensure there was no confusion about our AR simulation. However, we gave them *no* explanation about how to interact with virtual objects.

We then introduced participants to the task: moving a virtual apple to a nearby target. The target was represented by a circle on the virtual floor (Figure 7). Participants were told this task would have to be done “with the cube”. They were also specifically told to ignore the tablet for manipulation. Finally, we asked them to discover how to do that “by themselves, as far as possible”.

Obviously, we expected that some participants would not be able to complete the task without explanation. In order to understand how much help these participants needed, we designed a set of textual hints in increasing level of accuracy. During the task, there was a button that could be pressed to reveal a new hint. Each press on the button uncovered an additional hint on the tablet’s screen, in the following order:

1. “Put the cube onto the apple”
2. “Press the cube to grab the apple”
3. “Move the cube while maintaining the pressure”

These hints were specifically chosen to cover the different steps of object manipulation in our interface. Additionally, we chose to use textual hints rather than visual representations (e.g. arrows) to avoid ambiguous interpretations. We explained the role of the hint button to participants, and strongly encouraged them to use as few hints as possible.

## 5.3 Results and discussion

Figure 8 shows the percentage of participants according to the number of hints they needed to complete the task. A total of 19 participants (53%) successfully completed the task without requesting any hint. Among the remaining participants, all but one (45%) succeeded with the first two hints. The third hint was never used. We also report the task completion times for the participants who accomplished the task without any hint. The mean completion time was 63.5 s (SD=34.3 s).

More than half of participants discovered by themselves how to grasp and manipulate virtual objects with our tangible volume.

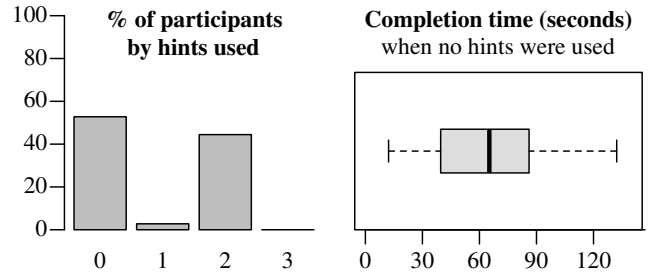


Figure 8: Number of hints and time needed by participants to discover how to manipulate a virtual object with our interface.

This is an encouraging result, given that our interface is so much different from the way most people currently interact with virtual 3D objects, and more generally interact with computers. At first glance, the completion times—about one minute on average—might seem long. However, this was the time needed to discover how to use a completely unfamiliar device *and* accomplish a manipulation task for the first time. This, it seems that the idea of grasping a virtual object “through” a tangible volume was spontaneously considered by a majority of participants, and within a reasonable time.

For those who did not complete the task without help, the number of hints requested provides more insight into which parts of manipulation were troublesome. The first hint was designed to uncover potential difficulties in positioning the tangible volume onto the virtual object. Nearly all participants who requested hints were not helped with this first hint. Hence, positioning was likely not what prevented them from completing the task. Indeed, all participants quickly noticed that a outline appeared around the virtual object whenever it was inside the volume. Some of them, however, believed that the outline meant the object was already selected, and attempted to move it without pressing the fingers.

Among other attempted strategies, a surprising number of participants attempted to “push” the virtual object with the tangible volume. This may indicate they thought the sides of the volume would be solid in the virtual scene. Many tried to tap or flick their fingers onto the surface of the volume. In some cases, it was apparently an attempt to replicate the “click” metaphor, especially when they grasped the tangible volume as if it were a mouse and tried to click (and even double-click) on the top. In other cases, it was clearly an attempt to affect the virtual object through the volume, especially when flicking a finger against one of the side faces. Finally, some participants tried to fully enclose the cube in their hands. This may have been an attempt to grasp the virtual object. However, in doing so they lost visual feedback, and it was thus impossible to see if finger pressure was sufficient to trigger the grasping technique.

Nearly all participants who requested the first hint also requested the second hint to complete the task. The second hint was designed to uncover difficulties with the grasping metaphor itself. No participant ever asked for the third hint. For those who needed help, the grasping metaphor was therefore the main hurdle. The third hint was about moving the selected object by keeping it attached to the tangible volume. Since no participant requested this hint, none of them encountered any problem with this last step.

## 6 USER STUDY: FIELD OF VIEW

Another challenge posed by our concept of a tangible volume is the limited field of view during manipulation. The tangible volume serves both as a display and as a tangible handle. Since our goal is to have a self-contained interface, the only display surface available is the surface of the volume itself. When the tangible volume is moved onto a virtual object to initiate manipulation, it also moves

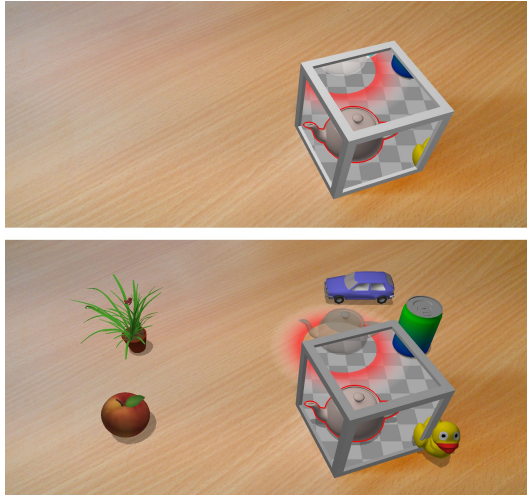


Figure 9: The two conditions in our second experiment. Top: the normal condition, with the view limited to the tangible volume (*narrow FoV*). Bottom: the control condition, with the view extending outside the volume (*wide FoV*).

further away from the user’s eyes, and its apparent size becomes smaller. Because of fish-tank rendering, the field of view on the virtual scene also becomes smaller. In contrast, moving the volume closer to the eyes would enlarge the field of view, but would also preclude direct manipulation.

In a study on handheld VR with mobile devices, Hwang et al. [12] observed that the motion of the device in space was sufficient to compensate for a limited field of view. However, in our case the tangible volume is smaller than many typical mobile devices, and is held further away from the eyes during manipulation. Overall, the average field of view with our tangible volume is much smaller than in a typical mobile device configuration.

If this limited field of view had a significant impact on spatial awareness, this could be an disadvantage of our approach of having the display area restricted to the tangible volume itself. On the other hand, it is possible that the motion of the device in space and the fish-tank effect still compensate for this limitation in our case. We thus conducted a second user study to measure the impact of the limited field of view on spatial awareness.

### 6.1 Participants and apparatus

This study was conducted with 32 participants. All of them took part in our previous user study, and thus had prior experience with object manipulation using our interface. The apparatus was still our AR-based prototype, with the pressure threshold restored to normal.

### 6.2 Procedure

The task was to navigate a virtual scene consisting of 6 virtual objects, and to search for targets that would appear after a delay. Each successive target contained a silhouette of a given virtual object (Figure 9), and this object had to be brought inside the target. After three of these manipulation tasks, the experiment stopped. Participants were then asked to indicate where they thought each virtual object was, by marking their location with small printed images. The goal was to evaluate how well they remembered the virtual scene, by comparing the reported object positions to the actual positions of virtual objects at the end of the experiment. The task itself was meant to simulate a typical use of the interface, by alternating navigation phases and manipulation phases.

The number of virtual objects in the scene (6) was selected to be high enough for evaluating spatial memory, but not too large which

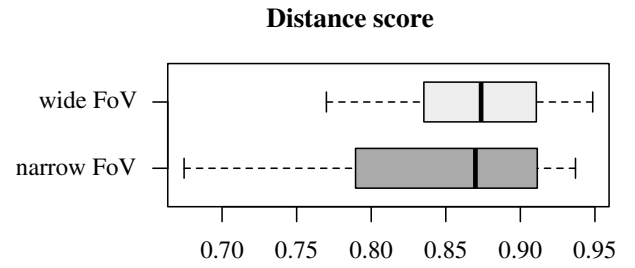


Figure 10: Correspondence between the positions reported from memory and the actual positions of virtual objects, based on the distance metric described by Sharlin et al. [25].

would make the scene too complex. The objects themselves were chosen to be easy to recognize and remember. The objects started in seemingly random—but predefined—positions in the virtual scene. Three targets appeared during the experiment, for three different virtual objects. Thus, some objects were manipulated and some remained untouched, simulating a typical manipulation workflow. Targets appeared in seemingly random, but actually predefined locations. After 15 s, a first target appeared. When the first manipulation was complete, a second target appeared after 10 s. After the second manipulation step, a third target appeared after 12 s. The third manipulation step was followed by a final delay of 20 s. Then, the display went blank and participants were invited to report the last positions of virtual objects. To do so, they were given small pictures of the 6 virtual objects, and asked to position the pictures in front of them where they thought the virtual objects were located before.

In order to evaluate spatial awareness with the tangible volume, we compared it to an artificial situation in which virtual objects could be seen *outside* the volume (Figure 9). In this control condition, we simply modified the rendering pipeline in our AR-based prototype so that virtual objects were also rendered outside the tangible volume. Approximately half of participants (17) performed the task in the *narrow FoV* condition, in which the field of view (FoV) was limited to the tangible volume. The others (15) performed the task in the *wide FoV* control condition, in which they could see the virtual objects outside the tangible volume during the entire task. Two additional participants in the control condition were removed due to unusable answers.

We used the *Distance* metric proposed by Sharlin et al. [25] to quantify the difference between the positions reported by participants and the actual positions of virtual objects.

### 6.3 Results and discussion

Figure 10 shows the distance scores computed from the participants’ answers, grouped by condition. The mean score in the *narrow FoV* condition (view limited to the tangible volume) was 0.845 (SD=0.080). The mean score in the *wide FoV* control condition (view extending outside the tangible volume) was 0.869 (SD=0.057).

Although there were a few lower scores in the *narrow FoV* condition, the overall results did not seem to differ much. To confirm this, we used a two sample t-test to compare the means of each condition. The results ( $t(30)=0.99$ ,  $p=0.32$ ) yielded a difference of means in the range  $[-0.075, 0.026]$  (95% CI). Thus, no practical difference between the scores of each condition was measured in our experiment.

From this result, it appears that the limited field of view in the *narrow* condition was not an important disadvantage compared to the control condition. This suggests that having visual feedback limited to the surface of the tangible volume is not a disadvantage for manipulation. This is an important point for portability, because

it eliminates the need for an external display surface.

This result also appears to confirm what was previously observed on mobile devices: moving a small display within the virtual scene may compensate for the limited field of view. Interestingly, in the narrow FoV condition, we noticed that some participants did bring the tangible volume closer in order to obtain a wider field of view. In a fully implemented tangible volume, with actual screens and head tracking, this would be a natural action to do. However, we were surprised to see that the tablet, which acted as the user's eyes in our prototype, did not discourage some participants from using this strategy.

## 7 BEYOND MANIPULATION: A TANGIBLE FRONTIER

We presented the advantages of turning a tangible object into a “volume of space” for creating a portable and self-contained TUI. We also demonstrated its use for 3D object manipulation, i.e. translating and rotating virtual objects. But the concept of a tangible volume has more potential than just for manipulation.

The real world and the virtual world are inherently separated. No object or information can directly move between them. Interfaces are used to transfer information from one side to the other. For example, most interfaces are designed to transfer user input coming from the real world, and visual feedback coming from the virtual world. Thus, the physical medium of the interface acts as a *frontier* that transfers information between the real and virtual worlds.

On a typical desktop computer, one such frontier is the monitor. The monitor constitutes a flat window that transfers visual information from the virtual world to the real world. However, user input comes from a completely separate location, usually the mouse and keyboard. With most current mobile devices, the frontier takes the form of a flat sheet of glass. This surface allows bidirectional transfer of information: light coming from virtual objects is transferred to the real world, and touch input on the surface is transferred to the virtual world. Since input and output happen on the same surface, the interaction is more direct. A mobile device is also portable: it can be freely moved in space to change the position of the frontier within both the real and virtual worlds. However, this frontier remains limited to a single planar surface, even though it is supposed to be an interface between two 3D worlds.

In our concept, we extend this idea by giving the frontier a volumetric shape. The tangible volume represents a volumetric frontier between the real and virtual worlds. Its surface provides a physical medium on which sensors and actuators can be attached, in order to transfer information between the real and virtual worlds. Since this tangible surface can be freely moved within both the real and virtual space, *every point of space becomes a potential interface* between the two worlds.

We demonstrated in the previous sections how this volumetric frontier can transfer information between the real and virtual worlds, in both directions. By attaching screens—a form of *actuators*—to the tangible volume and using fish-tank rendering, light from the virtual scene appears to exit the virtual world, and continues its path into the real world toward the user's eyes. By attaching pressure *sensors* to the tangible volume, force exerted by the user in the real world is directly transferred into the virtual scene.

Although we initially implemented these two modalities to support object manipulation, there is no reason why other modalities could not be transferred through the frontier as well. For example, light from the real world could be transferred to the virtual world, by attaching light sensors to the surface of the tangible volume. This would make it possible to render realistic lighting and shadows in the virtual scene (Figure 11(a)). Similarly, sound can be transferred in both directions by integrating microphones and speakers into the tangible volume. Another sense that can be exploited is the haptic modality. Forces originating from the virtual object could be transferred to the real world. By adding vibrators to the tangible volume,

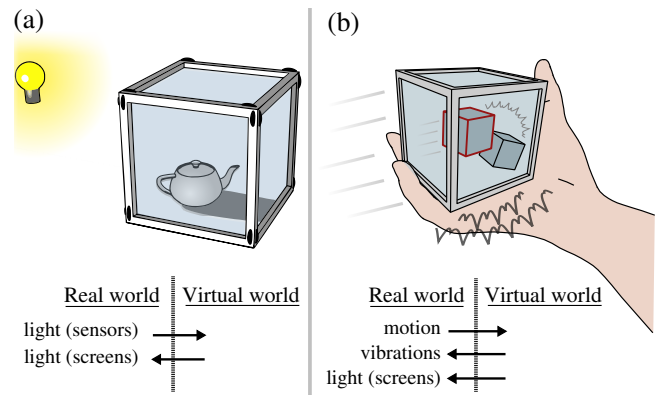


Figure 11: The tangible volume can be seen as a handheld “frontier” that transfers information between the real and virtual words. (a) by adding light sensors, light information can be transferred in both directions; (b) by adding vibrators, haptic feedback (e.g. collisions) from the currently selected virtual object can be transferred as well.

collisions between virtual objects could produce haptic feedback directly on the user's hand (Figure 11(b)). All of these forms of interaction can be implemented in a fully portable and self-contained interface, since the tangible volume provides the physical medium to attach any sensors and actuators. The shape of the frontier itself constitutes another possible interaction modality. Up to now, we only considered the case where the volume has a fixed shape (in our prototype: a cube). However, this shape could be made dynamic. By using flexible screens combined with underlying actuators, we can envision that this shape would change according to which virtual objects are located inside. This would bring the tangible frontier even closer to the virtual objects.

Examples where this concept of a tangible frontier could be used in practice include the interactive exploration of 3D data, computer gaming, and virtual prototyping. For scientific visualization, the tangible volume can provide a direct access to the three-dimensional data for probing, selection, and data manipulation through the freely movable frontier. In gaming, innovative game concepts would become available by transferring multiple modalities (light, sound, haptics...) in both directions through the surface of the volume, increasing immersion and user engagement. Finally, the fact that tangible manipulation and other forms of 3D interaction are offered by a single, portable and self-contained device can increase the opportunities for virtual prototyping. With the tangible volume alone, one could design a future product in virtual space and experience the sensation of holding it directly in the hand—with visual and haptic feedback. Furthermore, it would be possible to share the same experience with other designers or prospective customers by simply bringing along the tangible volume.

## 8 CONCLUSION

We introduced the concept of a “tangible volume”: a fully portable and self-contained device for 3D interaction, made of a single tangible object entirely covered with screens. This object represents a volume of the virtual scene, and can be positioned directly onto virtual objects. This makes 3D interaction more direct than with previous approaches. We described an object manipulation technique that consists in grasping virtual objects “through” the volume and moving them in 3D space. We created a partial prototype based on this concept, and used this prototype to investigate several usability questions. Future work could focus on three aspects. On the technical side, the first step would be to create a full implementation of our concept. This would require several improvements to the current technology. In particular, achieving reliable head tracking

and environment tracking in a fully portable and self-contained device remains an important challenge. On the experimental side, a more complete implementation would allow to repeat user studies in conditions that more closely resemble a real device. For example, the influence of correct finger occlusion could be studied in a prototype with actual screens, as well as the effect of possible imperfections in tracking and rendering. Another question to investigate is whether adding stereo rendering could help the user position the volume within the virtual scene. Finally, our concept itself can be extended to other forms of interaction. We demonstrated how our tangible volume can be used for 3D object manipulation. However, the concept of a tangible volume may be generalized as a volumetric “frontier” that can be positioned directly onto virtual objects. Future work could focus on exploring the full potential of this approach for other types of 3D interaction with a virtual world.

## REFERENCES

- [1] D. Avrahami, J. O. Wobbrock, and S. Izadi. Portico: tangible interaction on and around a tablet. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, page 347–356, 2011.
- [2] D. Baricevic, C. Lee, M. Turk, T. Hollerer, and D. Bowman. A hand-held AR magic lens with user-perspective rendering. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 197–206, 2012.
- [3] D. A. Bowman, E. Kruijff, J. J. LaViola Jr., and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, July 2004.
- [4] L. D. Brown, H. Hua, and C. Gao. A widget framework for augmented interaction in SCAPE. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*, UIST ’03, page 1–10, New York, NY, USA, 2003. ACM.
- [5] F. Ferreira, M. Cabral, O. Belloc, G. Miller, C. Kurashima, R. d. D. Lopes, I. Stavness, J. Anacleto, M. Zuffo, and S. Fels. Spheree: A 3D perspective-corrected interactive spherical scalable display. In *Proceedings of the 41st Conference on Computer Graphics and Interactive Techniques; Posters*, SIGGRAPH’14, page 86:1–86:1, Vancouver, Canada, Aug. 2014. ACM.
- [6] G. W. Fitzmaurice. *Graspable user interfaces*. PhD thesis, University of Toronto, Canada, 1996.
- [7] J. Francine and L. Nigay. Using the user’s point of view for interaction on mobile devices. In *23rd French Speaking Conference on Human-Computer Interaction*, IHM ’11, page 4:1–4:8, New York, NY, USA, 2011. ACM.
- [8] P. Harish and P. Narayanan. Designing perspectively correct multi-planar displays. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):407–419, Mar. 2013.
- [9] A. Henrysson, M. Billinghurst, and M. Ollila. Virtual object manipulation using a mobile phone. In *Proceedings of the 2005 international conference on Augmented tele-existence*, page 164–171, 2005.
- [10] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’94, page 452–458, New York, NY, USA, 1994. ACM.
- [11] K. Hinckley, J. Tullio, R. Pausch, D. Proffitt, and N. Kassell. Usability analysis of 3D rotation techniques. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, UIST ’97, page 1–10, New York, NY, USA, 1997. ACM.
- [12] J. Hwang, J. Jung, and G. J. Kim. Hand-held virtual reality: A feasibility study. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST ’06, page 356–363, New York, NY, USA, 2006. ACM.
- [13] M. Inami. Media3: The virtual hologram. In *ACM SIGGRAPH 97 Visual Proceedings: The Art and Interdisciplinary Programs of SIGGRAPH ’97*, SIGGRAPH ’97, page 107–, New York, NY, USA, 1997. ACM.
- [14] H. Ishii. The tangible user interface and its evolution. *Commun. ACM*, 51(6):32–36, 2008.
- [15] H. Ishii and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, page 234–241, 1997.
- [16] P. Issartel, F. Guénat, and M. Ammi. A portable interface for tangible exploration of volumetric data. In *Proceedings of the ACM 20th Symposium on Virtual Reality Software and Technology*, VRST ’14, page 209–210. ACM, Nov. 2014.
- [17] B. Jackson, T. Y. Lau, D. Schroeder, K. C. Toussaint, and D. F. Keefe. A lightweight tangible 3D interface for interactive visualization of thin fiber structures. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2802–2809, 2013.
- [18] R.-H. Liang. Augmenting the input space of portable displays using add-on hall-sensor grid. In *Proceedings of the adjunct publication of the 26th annual ACM symposium on User interface software and technology*, UIST ’13 Adjunct, page 33–36, New York, NY, USA, 2013. ACM.
- [19] R. Lopez-Gulliver, S. Yoshida, S. Yano, and N. Inoue. gCubik: real-time integral image rendering for a cubic 3D display. In *ACM SIGGRAPH 2009 Emerging Technologies*, SIGGRAPH ’09, page 11:1, New York, NY, USA, 2009. ACM.
- [20] A. Marzo, B. Bossavit, and M. Hachet. Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction*, SUI ’14, page 13–16, New York, NY, USA, 2014. ACM.
- [21] A. Mossel, B. Venditti, and H. Kaufmann. 3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality. In *Laval Virtual VRIC ’13*. ACM New York, NY, USA, 2013.
- [22] J. D. Mulder and R. Van Liere. The personal space station: Bringing interaction within reach. In *Proceedings of the virtual reality international conference, VRIC*, volume 2002, page 73–81, 2002.
- [23] D. A. Norman. *Things that make us smart: Defending human attributes in the age of the machine*. 1993.
- [24] W. Qi and J.-B. Martens. Tangible user interfaces for 3D clipping plane interaction with volumetric data: a case study. In *Proceedings of the 7th international conference on Multimodal interfaces*, ICMI ’05, page 252–258, New York, NY, USA, 2005. ACM.
- [25] E. Sharlin, B. Watson, S. Sutphen, L. Liu, R. Lederer, and J. Frazer. A tangible user interface for assessing cognitive mapping ability. *International Journal of Human-Computer Studies*, 67(3):269–278, 2009.
- [26] I. Stavness, B. Lam, and S. Fels. pCubee: A perspective-corrected handheld cubic display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 1381–1390, 2010.
- [27] I. Stavness, F. Vogt, and S. Fels. Cubee: A cubic 3D display for physics-based interaction. In *ACM SIGGRAPH 2006 Sketches*, SIGGRAPH ’06, New York, NY, USA, 2006. ACM.
- [28] C. Ware, K. Arthur, and K. S. Booth. Fish tank virtual reality. In *Proceedings of the INTERACT ’93 and CHI ’93 Conference on Human Factors in Computing Systems*, CHI ’93, page 37–42, New York, NY, USA, 1993. ACM.