Tangible VR Book: Exploring the Design Space of Marker-Based Tangible Interfaces for Virtual Reality

Jorge C. S. Cardoso [1,*,†] and Jorge M. Ribeiro [2,†]

1 Centre for Informatics and Systems of the University of Coimbra, Department of Informatics Engineering, University of Coimbra, 3030-790 Coimbra, Portugal
2 Department of Informatics Engineering, University of Coimbra, 3030-790 Coimbra, Portugal; jorgeribeiro@uc.pt
* Correspondence: jorgecardoso@dei.uc.pt
† These authors contributed equally to this work.

Abstract: Tangible User Interface (TUI) represents a huge potential for Virtual Reality (VR) because tangibles can naturally provide rich haptic cues which are often missing in VR experiences that make use of standard controllers. We are particularly interested in implementing TUIs for smartphone-based VR, given the lower usage barrier and easy deployment. In order to keep the overall system simple and accessible, we have explored object detection through visual markers, using the smartphone's camera. In order to help VR experience designers, in this work we present a design space for marker-based TUI for VR. We have mapped this design space by developing several marker-based tangible interaction prototypes and through a formative study with professionals with different backgrounds. We then instantiated the design space in a Tangible VR Book which we evaluate with remote user studies inspired by the vignette methodology.

Keywords: virtual reality; smartphone-based virtual reality; tangible user interface; tangible book

1. Introduction

Ishii and Ullmer [1] introduced the concept of TUIs more than 20 years ago, and later systematised and defined it as:

> ... tangible interfaces give physical form to digital information, employing physical artifacts both as representations and controls for computational media [2].

For VR, tangible interaction represents a huge potential because physical objects can naturally provide rich haptic cues which are often missing in VR experiences. Users often manipulate virtual objects that have clear physical counterparts, through generic controllers in VR, experiencing the exact same sense of weight, temperature, texture, etc., no matter the virtual object. TUIs can provide more natural interactions [3] and would correspond to a better VR solution according to Carrozzino and Bergamasco [4]. TUIs can also result in higher immersiveness [5,6] because they provide richer sensory feedback than standard controllers. They can also contribute to higher feeling of presence, particularly physical presence as defined by Lee [7]. At the very least, TUIs can be more fun and engaging [8].

Implementing tangible interaction within a Virtual Environment (VE) requires the system to be able to detect physical objects, ideally with positional and orientation tracking—six Degrees of Freedom (DOF). Over the years, various alternatives have been tried, from active (instrumented with sensors) to passive objects.

In our work, we leverage the mature marker recognition algorithms for smartphones as a means to detect physical objects while the user is immersed in a VE. This approach has the advantage of being applicable to smartphone-based VR and making use of already existing hardware on the smartphone: the camera. In addition, visual markers are easy to print and fit to various kinds of physical objects, making this approach very cheap and accessible. These characteristics make this approach suitable for Do It Yourself (DIY)

scenarios where users receive a template or a kit to assemble but also to in-class situations where a large number of tangibles and VR systems are needed. Because TUIs are also generally equated with ease of use, our approach is also suitable for walk-up-and-use scenarios where users do not have much time to learn how to use the system such as museums and other cultural heritage exploration situations.

Although there have been some applications to VR, e.g., [9], the design space for marker-based TUIs for VR has not been sufficiently well studied.

Our primary contribution is the definition and exploration of a design space of marker-based TUI for VR. We first adopted a research-through-design inspired process and produced several prototypes to help us gain insight about the possibilities of marker-based TUIs. Then, similarly to Surale et al. [10], we performed a formative study where we asked participants to envision possible uses and expectations for marker-based TUIs for VR. Based on the proposed ideas and on our own exploration we mapped out a design space with eight dimensions that can help designers understand alternatives for designing marker-based TUIs. We instantiated the design space in a Tangible VR Book for architectural explorations that we then evaluated.

2. Background and Related Work

This section provides an overview of how tangibles have been used in VR, what kind of marker-based TUIs exist outside of VR, and on interfaces that take the form factor of a book—the same form factor we chose for our Tangible VR Book.

2.1. Tangibles in VR

In general, tangibles can be implemented in two very different ways: they can be active objects that require power to operate, or passive objects.

2.1.1. Active

Johnson et al. [11] implemented plush toys with various sensors including pitch and roll sensors, gyroscope, and magnetometers as well flexion and squeeze sensors, which enabled them to sense various kinds of interactions with the toys. Although the usage of the toys was not in immersive VR but rather in front of a screen, this project demonstrates an interesting aspect of adapting existing toys: they are familiar to users and they can easily be "diegetic tangibles" that "exist within the space and time of a narrative's world and can be an effective strategy for interaction design and for narrative design" [12].

Sajjadi et al. [13] employed Sifteo Cubes as tangible objects in a VR game—Maze Commander—where two players, one using a VR headset and another manipulating the Sifteo Cubes, collaborate to escape a maze. The player with the Sifteo Cubes, moves them around to form and manipulate a virtual maze, while the VR user can identify enemies and obstacles. Sifteo Cubes are small computers (cubes with about 4 cm) with screens and sensors that enable them to react to movements and proximity with one another. Maze Commander demonstrates an important aspect of tangible interaction: part of the state of the VR system is visible and physically manipulable by users outside of VR.

Another example is Snake Charmer by Araujo et al. [6] where a robotic arm acts as physical object to provide haptic feedback. The arm tracks the movements of the user's hand and is able to reposition itself so as to provide haptic feedback when the user touches a virtual object. It is also capable of picking up different endpoints to provide different surface textures, temperatures, etc. Snake Charmer highlights the importance of the detail of the haptic feedback when interacting in immersive VR.

All these are examples of active tangibles: objects instrumented with sensing capabilities that need power to operate. Active tangibles are usually more expensive and/or require considerable effort to produce.

2.1.2. Passive

On the other hand, even simple passive tangibles have been shown to have a significant positive effect on the VR experience [5].

Aguerreche et al. [14], for example, created a reconfigurable object with the shape of a triangle with extendable edges. This passive object is detected by an external sensing infrastructure (motion capture studio) and can be associated with various kinds of virtual objects, allowing their manipulation.

Passive objects can also be detected by external depth sensing cameras. In the Annexing Reality system [15], for example, a Kinect sensor is use to identify physical objects and map them to virtual objects with similar shape. Users can then pick and inspect the virtual object while actually picking up and manipulating a similar physical object. Depth cameras however, are not yet found in mainstream devices and the Annexing Reality prototype with a Kinect sensor was cumbersome to use.

These two examples highlight the importance of flexible solutions where the same object can represent different virtual ones and opportunistic use of everyday objects so that creating tangibles can be cheap and rapid.

2.2. Marker-Based Tangible User Interfaces

Although Visual markers are mostly associated with Augmented Reality (AR) experiences, they have been used extensively for various kinds of TUIs. Jord et al. [16] used fiducial markers in their Reactable system for music performance. Zheng et al. [17] demonstrates how various types of TUI prototypes can be created with printed paper markers. Ikeda and Tsukada [18] combine visual markers with capacitive sensing to create tangible that can be hovered over multitouch capacitive displays. Conversely, there are not many examples of visual markers used in immersive VR. Although not strictly for VR, the following examples demonstrate the possibles uses of visual markers for detecting objects and provide good examples of how they could be used in smartphone-based VR.

Henderson and Feiner [19] created a class of interaction techniques for AR which they called opportunistic controls: "a tangible user interface [Ishii and Ullmer 1997] that leverages naturally occurring, tactilely interesting, and otherwise unused affordances" [19]. In their implementation, they use structured visual markers to compute the position and orientation of the natural physical objects. Additional computer vision algorithms can then detect gestures over the regions of interest relative to the visual markers. In our Tangible VR Book prototype we rely instead on the marker tracking functionality to detect simple gestures over the markers themselves by tracking the sequence of markers that become hidden/visible.

Paolis et al. [20] implemented a billiards simulation using marker detection. Markers were placed on a surface to provide a reference for the billiards table and on the tip of a physical cue. Although the visualization of the simulation was on a desktop display, this example is a further demonstration of how versatile visual markers can be and how easy it is to use them to track physical objects. In a similar vein, Cheng et al. [21], in their iCon system, stuck fiducial markers on everyday objects to convert them into input controllers for various applications.

Lee et al. [22] provide an interesting example of visual markers applied in gesture-based tangible interactions for mixed-reality environments. In this case, visual markers are used not to track an everyday object, but a part of a user's body: the hand. In our prototype for the Tangible VR Book, we also experimented with visual markers on the user's hands as a way to provide visual feedback about the position of the hands.

2.3. Book-Based Interfaces

We instantiate the design space for marker-based TUIs with the Tangible VR Book. It is thus relevant to analyse other projects that have used a book form as the interface.

The Magic Book [23] is an "enhanced version of a traditional pop-up book", that allows users to scan AR markers on the pages to see 3D virtual models. Users can quickly

switch to an immersive mode and enter the 3D model. The book is experienced through a handheld AR display, so users do not manipulate the book continuously, but rather lay it down on a table and look at it through the AR display. While inside the 3D model, the physical book object is not needed. The book is meant as a collaborative interface that allows quick transition between immersive VR and AR modes. Our Tangible VR Book employs the same object detection technology (visual markers) but is intended for manipulation while inside an immersive VE. We also employ a number of interactive features not available in the Magic Book.

The Companion Novel [24] is an interactive device with a book-like appearance that visitors can take with them while visiting Sheffield General Cemetery to automatically hear audio stories associated with to specific locations and themes. The system is complemented by Bluetooth speakers close to the points of interest that detect the book's presence to render audio recordings. The book format was purposefully meant to provide intuitive handling as the interactions with the device are accomplished through familiar actions with regular books. Interaction with the system is essentially accomplished by opening the book and placing a bookmark on the page associated with the chosen theme. The book format of the Companion Novel was meant to "fit in with the environment in an unobtrusive way" [24]. Similarly, we chose the book format for our main prototype because it is an eclectic form factor that "belongs" anywhere. Contrary to the Companion Novel, our system is meant for usage within a VE and leverages on additional interactions to control the display of associated information.

Other interfaces that demonstrate the eclecticism of the book form factor are the Paper-Video [25], which investigates how interaction with video can leverage on the existence of paper-like displays in a book-like form factor, and FoldMe [26], which explores interaction with foldable displays through a book-like device. Unlike these projects, which aim at real world interactions with physical devices, our Tangible VR Book is specifically aimed at immersive VR usage. Although we have not followed this avenue, the design spaces defined by PaperVideo and FoldMe could be also applied to a TUIs similar to the Tangible VR Book.

2.4. TUI Frameworks

The use of physical objects to interact with digital systems has been studied and structured according to very different perspectives over time [27]. In this section we briefly review a few of the tangible interaction frameworks that have been proposed and how they relate to the design space we propose.

Holmquist et al. [28] classifies the possible associations between physical objects and digital information, according to the informational role that the objects take, as containers, tokens, and tools. Containers are generic objects that can be associated with any kind of digital information. For a container, the digital information associated with it can vary throughout the object's lifetime. Tokens are objects whose physical properties reflect the digital information associated with them. The information associated with a token typically does not change over time because it is closely tied to that particular physical object. Finally, tools are objects used to manipulate digital information.

Ullmer and Ishii [29] classified TUI according to how multiple objects are interpreted to form a system of physical objects. In spatial systems, the spatial configuration of the objects within some reference frame is used to define the state of the tangibles. In relational systems, "the sequence, adjacencies, or other logical relationships between systems of multiple tangibles are mapped to computational interpretations" [29]. Constructive systems represent tangibles that have modular elements that can be assembled together (like LEGO™ blocks, for example) to create different interpretations. Finally, associative tangibles represents tangibles that "do not reference other objects to derive meaning" [29].

Hornecker and Buur [30] proposed a "framework that focuses on the user experience of interaction and aims to unpack the interweaving of the material/physical and the social aspects of interaction". Their framework on physical space and social interaction for

tangible interaction is composed of four themes that provide different lenses to think about the design of tangible systems. Each theme is subsequently decomposed into concepts that represent analytical tools for describing and interpreting tangible systems. Tangible manipulation refers to the bodily interaction with physical objects that are somehow coupled with computational functions. This theme is concerned with the material qualities of the objects and how their manipulation is associated with the effects on the system. Concepts in tangible manipulation are: haptic direct manipulation (ability to grab, feel, and move), lightweight interaction (ability to interact in small steps with immediate feedback), isomorph effects (ability to understand the relation between actions and their effects). Spatial interaction refers to how tangible interaction is embedded in physical space and requires (bodily) movement in space. Spatial interaction includes concepts such as: inhabited space (whether people and objects meet and form a meaningful place), configurable materials (space is configurable and appropriatable, moving things around has meaning), non-fragmented visibility (whether everybody can see what is happening), full-body interaction (whether we can use our whole body), performative action (whether our body movements can communicate something while we are doing what we do). Embodied facilitation refers to how the physical configuration of objects in space directs group behaviour. Concepts in embodied facilitation are: embodied constraints (whether the physical configuration constrains behavior leading to collaboration), multiple access points (whether all users can see what is happening and participate by using the central objects), tailored representations (whether the representations are familiar and connect with users' skills). Expressive representation focuses on what tangibles represent, how expressive they are, what material and digital representations are employed, and how they are perceived by users.

Simeone et al. [31] define a layered model for Substitutional Reality [32] where a VE is adapted to the physical world by a process of substitution of physical objects with virtual ones. At each, of the five levels of substitution, the virtual object gets further away from its physical counterpart. In the replica level, the virtual object appears as a 3D model that replicates the physical proxy, including providing exactly the same affordances and allowing the same interactions. In the aesthetic level, physical properties of the appearance of the virtual object may be different from its physical counterpart. For example, a virtual object may be rendered with a different material, which in turn might affect how it is perceived by users. In the addition/subtraction level, virtual elements are added or removed from their physical counterpart objects. For example, a mug may be virtually depicted without a handle (subtraction). The function level represents substitutions where there is no longer a match between all the affordances [33] of the physical and virtual objects. For example, a "mug might be rendered as an oil lamp as they share some affordances (we can pour liquids in both) but are intended for two different functions (drinking or lighting)" [31]. Finally, the category level represents substitutions that reach the extreme where the appearance of the virtual object has no connection to its physical counterpart.

Harley et al. [34] describe a framework for tangible narratives that rests on seven categories. Each category "reflects narrative possibilities or constraints, shaping how the narrative is created and communicated". Primary user(s) represents the intended audience (children, teenagers, adults). Media represents how the feedback after tangible interaction is presented to the user. Narrative function of the tangible objects represents the use of the tangible in the narrative (character, navigation tool, metaphor for story component, etc.). Diegetic tangibles indicates whether the tangibles can be considered diegetic, i.e., the object exists within the narrative world. Narrative creation considers whether the system allows users to create and/or tell stories. Narrative choice indicates what type of choices users can make during the storytelling which may influence the narrative's path. Finally, narrative position represents the user's role within the narrative (external-exploratory, internal-exploratory, external-ontological, and internal-ontological).

3. Early Prototyping

The goal of the early prototyping phase was to provide us with insights and knowledge about the realization possibilities and limitations of marker-based TUIs. We adopted a research-through-design inspired process similar to the one by Versteeg et al. [35].

In this section we briefly describe some of the prototypes that we developed and the reflections that resulted. Figure 1 shows some of these early prototypes.
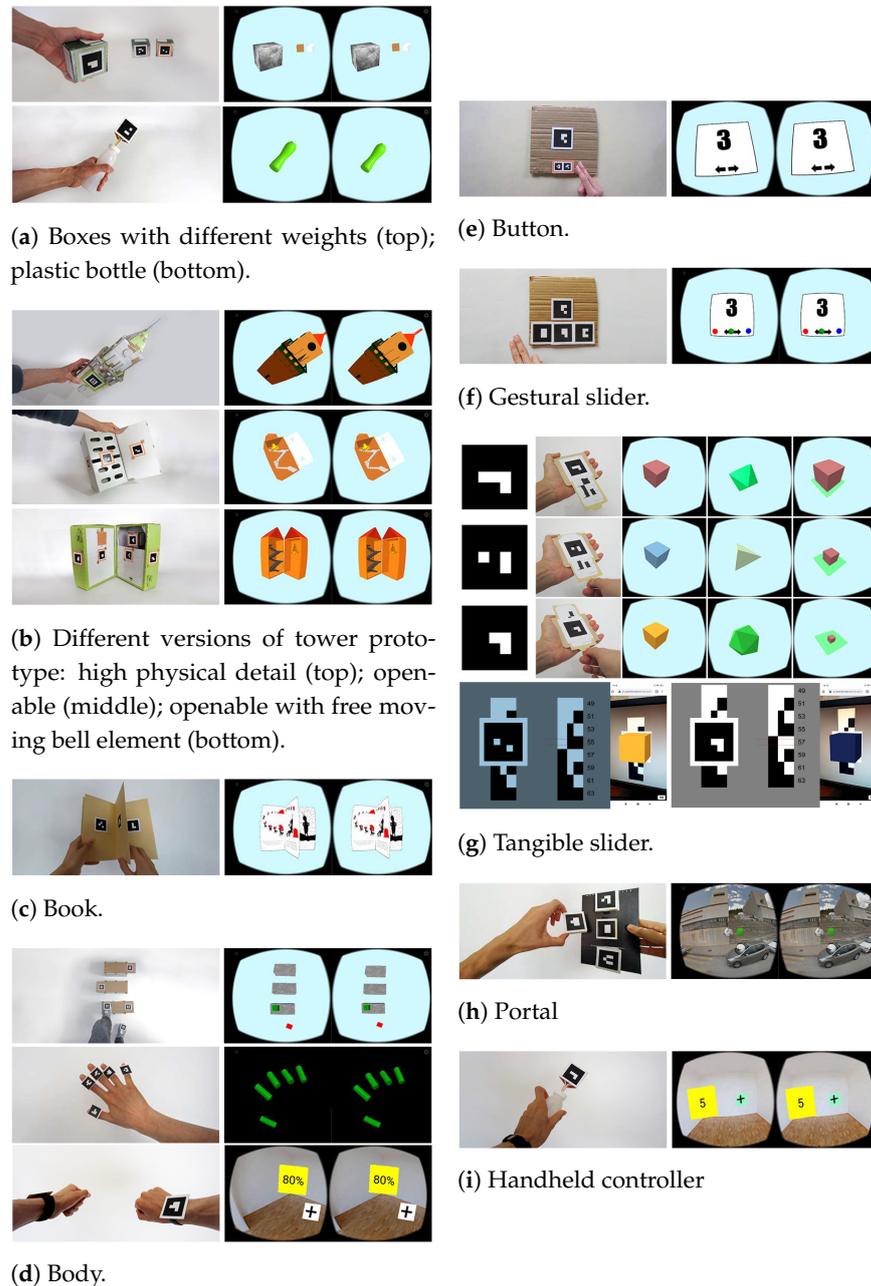

(**a**) Boxes with different weights (top); plastic bottle (bottom).

(**e**) Button.

(**f**) Gestural slider.

(**b**) Different versions of tower prototype: high physical detail (top); openable (middle); openable with free moving bell element (bottom).

(**c**) Book.

(**g**) Tangible slider.

(**h**) Portal

(**i**) Handheld controller

(**d**) Body.

Figure 1. Selection of early marker-based TUI for VR prototypes.

The initial prototypes (Figure 1a) were very simple and meant mostly for a basic understanding of how different physical properties could be simulated.

Additional prototypes started exploring more complex forms and manipulations. Figure 1b shows three different prototypes for a tangible tower. The top version explored the detail in the physical object so as to mimic the architectural detail of the tower. The middle and bottom versions allow users to "open up" the tangible and inspect its interior. The bottom version further includes a free-moving element to represent the tower's bell: as the tangible is moved

this element will naturally swing both in the physical and in the virtual environment. A book object (Figure 1c) was also prototyped. As users flip through the pages, they can see visual renderings of the page and simultaneously hear the audio version.

We also explored the use of markers in and in relation to the body. Figure 1d shows three of the prototyped ideas: the use of markers on the feet and on "staircase steps", on the fingers, and on the wrists. Markers on the body serve the main purpose of allowing the system to provide visual feedback about the user's body, but they could also allow simple gesture detection for input.

We also explored various kinds of input controllers such as "buttons", gestural and tangible sliders, portals, and handheld controllers (Figure 1e–i). The button controller works by detecting the absence/presence pattern over a specific marker (Figure 1e). The gestural slider works by detecting an absence/presence pattern over a succession of markers (Figure 1f). The tangible slider works by dynamically displaying different markers by sliding over a strip with several pre-coded marker patterns (Figure 1g). The portal works by detaching a markers from a a surface and bringing it closer than a pre-defined threshold to the user's eyes (Figure 1h). Finally, the handheld controller makes use of the earlier physical bottle prototype and detects proximity to other objects or rotation of the controller itself through the succession of visible markers (Figure 1i).

Reflections

Some physical properties are easy to convey such as form, size, weight, and texture even with cheap materials. Other properties such as smell or acoustics are harder to imitate.

Markers are hard to include in the surface of some objects such as the plastic bottle because they need to be on flat surfaces. In these cases, the physical object may need to be adapted in order to allow or facilitate tracking such as extending it beyond it natural limits (Figure 1a).

In our prototypes, we used a standard smartphone without any special camera lenses. This means that the field of vision is limited and particularly noticeable in tangibles made up from several parts as the openable tower. Without enough markers in the Field of View (FOV), some parts of the virtual representation suddenly disappear. Despite these issues, the effect of holding a physical architectural mockup and being able to bring it closely to our eyes, opening it up to look inside and inspect minute details is very compelling.

The book prototype resulted in an easy to manipulate tangible where the natural affordances of a book are mostly directly transported to the VE. We can easily imagine taking advantage of the richness of digital information while keeping the manipulation simplicity of the book. In addition, the form factor results in an inconspicuous object that "belongs" most anywhere without attracting attention to its use.

A shortcoming of the use of markers on the feet is that the markers need to be large in order to be recognisable, but this makes them less wearable. The use of markers on the hands, particularly in the context of some sort of wearable device, may be a cheap solution for hand detection on smartphone-based VR.

Some types of controllers such as the button, gestural slider, or handheld controller are hard to implement in a stable manner. It is easy to unintended activations due to the markers disappearing from the camera's view. The tangible slider and detachable portals do not suffer from this problem. The tangible slider resulted in an interesting way to create dynamic objects where the associated marker can be dynamically changed by moving the underlying strip. Additionally, it suggests many other control variations, for example one where the strip automatically goes back to a resting position when the control is not being manipulated.

4. Formative Study

The goal of the formative study was to understand how people envision using marker-based TUIs in VR. Contrary to [10], we did not have a specific application in mind. Additionally, we considered that it would be difficult for participants to imagine something (interacting through physical objects in VR) that would be very far from their experience. Instead of asking participants to act out specific tasks, we allowed them to experience

simple TUI VR prototypes and then interviewed them. The results from these interviews were later used to build a design space for marker-based TUI for VR.

4.1. Participants

Fourteen participants among the university staff and students were recruited through email, or opportunistically through direct contact on several locations. Participants were from several backgrounds/professions: civil engineering, administrative worker, web designer, theology professor, IT technician, tourism, sales, informatics engineering, architecture, pharmacy. Ages ranged from 27 to 55 years old, but most (57%) were aged between (40–49) years old. Six participants were female and eight were male. Four participants reported no previous experience with VR, six reported basic experience, and four reported regular use, but no one reported being expert in VR.

4.2. Procedure

Sessions were carried out in our lab, or at the participants' workplaces. Participants were introduced to the purpose of the study and filled in a personal data questionnaire. They were then asked to try out three prototypes using a smartphone-based VR system provided by us. Although the prototypes were very simple, we guided participants' experimentation in order to make sure that all relevant aspects were experienced. Participants were then interviewed to gather their feedback and ideas regarding the tangible interaction paradigm. During their interaction with the prototypes, participants were encouraged to externalise their thoughts and specific questions were asked to promote reflection. The interview was semi-structured with a few planned questions regarding what contexts and applications participants could imagine given the prototypes they had just explored, what could be improved in the various prototypes, what other ideas did the prototypes suggest. The main purpose of these questions and of the prototypes was to engage the participants' imagination in order to elicit further ideas that could feed the design space. Data from the interviews were audio-recorded and later transcribed. Some participants contacted us by email after the sessions, providing additional ideas they had not occurred to them during the session.

4.3. Results

This section presents the main results from the formative study, coded with R# in order to facilitate referencing later in the paper.

R1. Physical characteristics of objects

Participants often mentioned specific characteristics of physical objects that might be relevant to convey in a TUI, "Be able to feel the texture and shapes of certain parts of a monument." [P9], "feel the smell of the lamp in the chapel" [P11]. Geometric shapes, textures, volume, density, smell are some of the characteristics mentioned by participants.

R2. Object for system input

Several users suggested the use objects as a way to directly send input to the system: "use the object to hit the tower's bell and reproduce the sound" [P7], "control to see a cut of the floor" [P13]. Participants suggested various kinds of input such as using the tangible to "hit" another object, to change location, to use it as a key to activate content, as a controller to change music.

R3. Output modality associated with object's manipulation

Participants also mentioned several examples of what they expected to see or hear as the result of the manipulation of the tangible. Interestingly, audio was an often mentioned output modality: "see the chapel's organ and be able to press the keys to reproduce the different sound of the tubes" [P11], ". . . book is accompanied by ambient sound of the respective eras" [P4], "the experience can be acompanied with audio of languages of other

eras such as archaic Portuguese, latin, etc." [P4]. Video and 3D models were also often mentioned, but animations, text, images were all also suggested.

R4. Object as replica

Tangibles were often associated with existing physical objects as a substitute to the experience of holding or manipulating the original physical object: "Materialize objects of religious art . . . because they represent something very important but cannot be available to the public" [P4], "Use the 'tightrope walker' of the physics department and we control it equilibrium position" [P10].

R5. Deconstructing objects

Participants also often described situations where it would be useful to be able to open up and deconstruct an object: "Open the computer to know what is inside and understand how the connected components work" [P2], ". . . can be opened to explore the interior of old buildings . . ." [P3], "Open a computer and use objects as components" [P5], "Disassemble a construction model by its parts to see inside" [P13].

R6. Magical interactions

Some of the ideas offered by participants referred interactions that went beyond simple physical manipulation, but instead represented magical interactions that are impossible in our physical world: "With the architectural model, I could choose a specific apartment so that clients could go inside the space" [P1].

R7. Fixed objects

Although most ideas referred to the manipulation of small objects, a few ideas referred to fixed/large objects existing in the environment: "Climb a few of the steps to see the area of the tower's bell" [P7].

5. Design Space

Combining the ideas resulting from the formative study and our prototype exploration, we shaped a design space consisting of eight dimensions that provide different design options (Figure 2). Not all results were directly applied in the design space. Conversely, some of the dimension of the design space were inspired from related work. We use the term "dimensions" as synonym to the "questions" of MacLean et al. [36] in their Design Space Analysis method. The eight dimensions are described in the following subsections.

Our design space does not attempt at describing the interactive elements, other more specialized frameworks exist for this, such as [37] design considerations on spatial input, Ref. [38] Token and Constraints paradigm for tangibles interfaces, or [39] framework on Expected, Sensed, and Desired movements of interfaces.

Figure 2. Dimensions and options of the design space.

The options in the each dimension of the design space are not meant to be mutually exclusive, even though we may want to highlight some aspects of the tangible by considering one option instead of another. The (partial) exception is on the type of object

dimension where options are mutually exclusive within the pair (movable/immovable, reconfigurable/non-reconfigurable, active/passive, dynamic/static).

5.1. Overall Fidelity

The overall fidelity dimension captures part of the layered model of substitutional environments by Simeone et al. [31]. We simplify the layered model however, by considering only two options: replica, and proxy.

As the name implies, a replica is a reproduction of an existing physical object. How faithful the replica should, must be subjectively defined by the design team but the intention is that users perceive the object as replicating an existing physical one. The replica option is sufficiently important to be maintained in our design space because it represents a recurrent reference in the formative study [R4]: the possibility of picking up and inspecting a known physical object that is otherwise not possible (because it is inaccessible, for example).

The proxy option includes the remaining layers of substitutional environments: aesthetic, addition/subtraction, function, category. In essence, a proxy physical object may be far away in shape, weight, size, etc., from the virtual object it represents but still provides users with a tangible element.

5.2. Main Physical Characteristic

The main physical characteristic dimension captures the focus of the intention when designing the tangible object for a particular VR experience. Maybe what is most important for a given experience is the shape of the object, regardless of its weight or texture; or perhaps it is the rigidity that the designer wishes to convey... [R1]. The options for this dimension (shape, rigidity, size, weight, texture, temperature, smell, acoustics) can serve as reminders to designers that objects have many non-visual characteristics that can be explored through tangibles. It is also a reminder that there will certainly be many characteristics that are not possible to convey, unless the tangible is the physical object being represented in VR. The acoustics option deserves a special mention: the acoustic properties of a physical object were only indirectly mentioned in the formative study sessions and we have not found any example of a TUI that explores acoustics in the literature. However, it is an option that follows from the other non-visual properties of objects that we can somehow explore haptically, e.g., by tapping with one's fingers on an object or by hitting an object with another. The list of options in this dimension is not necessarily complete, although we consider texture as a characteristic including most of the sub-senses of touch.

5.3. Main I/O Role

Although tangibles are always used both as input to the system [R2] and as representation of the state of the system, some tangibles may emphasise one over the other. The main I/O role dimension represents this, and allows designers to consider the most relevant aspects of the overall system. For example, if the tangible's main role is input then it will be used to somehow affect the state of the VE. In this case, designers need to consider how the tangible will be manipulated and how the system will detect the input, e.g., will the system detect gestures, relative position or orientation of the tangible, proximity to other tangibles or virtual objects, etc. Designers will also need to consider the effect of that manipulation on the VE. If the tangible's main role is output, designers need to consider which type of information will be associated with the tangible and whether that information will be visualised on the virtual object or somewhere else in the VE.

In some situations it may not be easy to decide if the main role of a tangible is to act as a controller (input) or as output of information. Some tangibles may simply act as a physical representation of a virtual object that users can manipulate to inspect the virtual object without affecting anything else in the VE. In many cases where the tangible's overall fidelity is a replica, its main I/O role will simply be undefined because it is acting as a replica for a passive, inert object. However, some tangibles may replicate tools, for example,

in which case they will be used primarily as input. In other situations, the tangible's role may be defined as input and output without one prevailing over the other.

5.4. Output Modality

The Output modality dimension refers to how the tangible is represented inside the VE and to the type of media associated with it (i.e., when the tangible's main I/O role is for output) [R3]. Most often the tangible is represented through a 3D model in the VE, but other options are possible: text, image, graphics (e.g., computer generated graphics or animations), audio, video. Considering the output modality is important because it has implications in other aspects of the interaction. If a tangible is associated with time-based media such as audio or video, designers need also to consider how the reproduction of the media will be controlled (manually by the user, or automatically by the system). Another aspect to consider is where the associated output information will be rendered. This is covered by the next dimension.

5.5. Output Coupling

The output coupling dimension refers to whether the information associated with the tangible is rendered as part of the virtual object itself or on the environment. We consider three options for this: the information is rendered on the object itself, as if being a part of the object; the information is coupled with the object (if the object moves in 3D space, the associated information will move accordingly) but is not rendered directly on the object; the information is rendered somewhere in the environment independently from the object.

5.6. Type of Object

The type of object dimension combines several binary properties that the physical object can exhibit [R5, R7]: movable/immovable, reconfigurable/non-reconfigurable, active/passive, dynamic/static.

Immovable objects are usually large and part of the architecture of the space (walls, doors, large furniture). Movable objects can be moved, picked up and manipulated. Movable objects generally need more consideration and attention to detail because they will most often be held in users hands for longer periods of time.

Reconfigurable objects have moving or deformable parts and can thus somehow change shape during their use. For example the openable tower prototype would be considered reconfigurable as it could be opened and closed. A non-reconfigurable tangible cannot change physical shape during its use.

Active tangibles depend on electrical current to function, while passive tangibles do not.

Dynamic tangibles are objects whose meaning in the VE can change over time, while static objects have the same virtual representation throughout the whole VR experience. It should be noted that all physical objects can be used as dynamic tangibles if the system is somehow able to assign different meanings throughout time. It is not absolutely necessary that the physical object is able to change configuration or state, although that is one solution. The tangible slider early prototype would be an example of a dynamic tangible.

5.7. System Configuration

The system configuration dimension is directly taken from and equivalent to the categories defined by Ullmer and Ishii [29]: spatial, relational, constructive, and associative.

The first three options represent systems of multiple interdependent tangibles. In spatial systems, the spatial configuration of the objects within some reference frame is used to define the state of the tangibles. In relational systems, "the sequence, adjacencies, or other logical relationships between systems of multiple tangibles are mapped to computational interpretations" [29]. Constructive systems represent tangibles that have modular elements that can be assembled together (like LEGO™ blocks, for example) to create different interpretations. Finally, the last option—associative—represents tangibles that "do not reference other objects to derive meaning" [29].

5.8. Informational Role

The informational role dimension is equivalent to Holmquist et al. [28] classes of physical objects that represent digital information: containers, tokens, and tools.

Containers are generic objects that can be associated with any kind of digital information. For a container, the digital information associated with it can vary throughout the object's lifetime. Tokens are objects whose physical properties reflect the digital information associated with them. The information associated with a token typically does not change over time because it is closely tied to that particular physical object. Finally, tools are objects used to manipulate digital information.

6. Tangible VR Book

The Tangible VR Book (Figure 3) is similar to a physical book with thick pages that contain visual markers on each page. This roughly keeps the same affordances [40] as those of a normal book, given the book can be picked up to inspect the front and back cover, to turn pages, to bring closer or farther away from one's eyes for reading. These affordances are the same for the virtual representation of the VR book. The Tangible VR Book does not attempt to be a replica of any specific book thus being a proxy object, in the overall fidelity dimension.
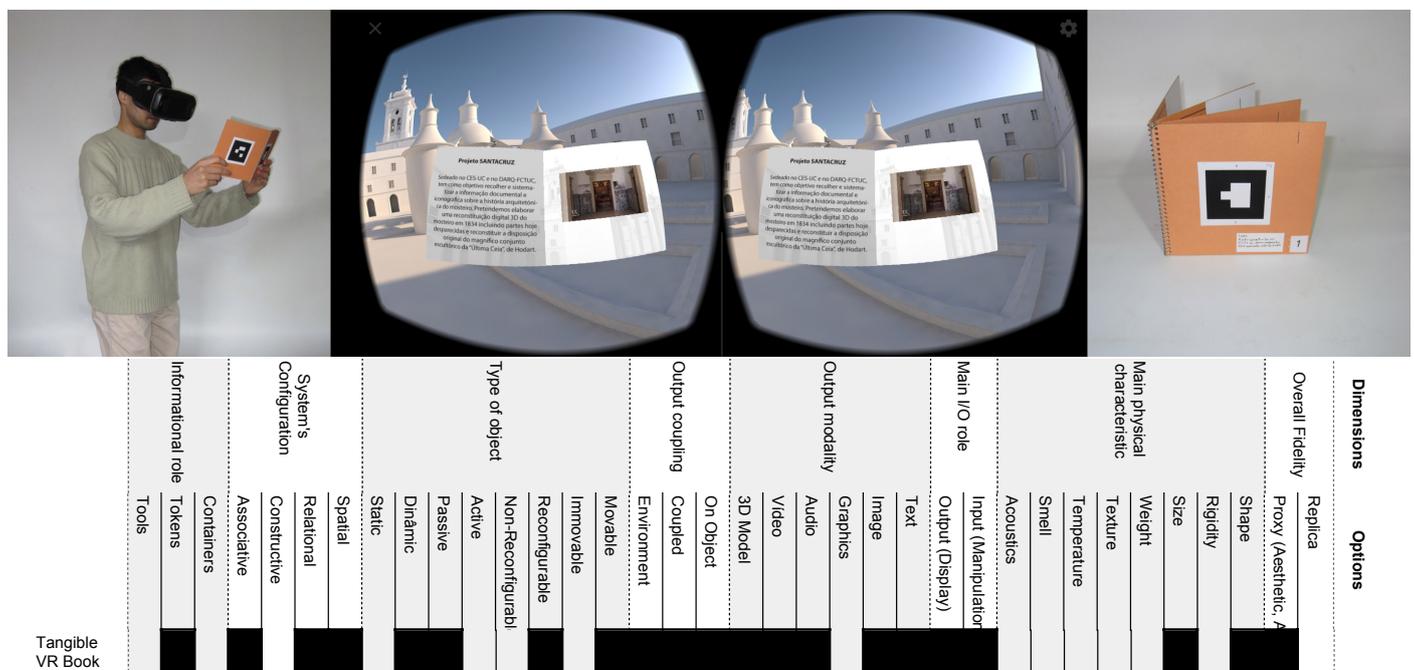
Figure 3. Overview of the Tangible VR Book (**top**); Instantiation of the Marker-based TUI Design Space in the Tangible VR Book (**bottom**).

We opted for the book format for several reasons. The book is a familiar object, that we learn to use from very young ages—there are books designed for children under one year of age. Books are also very eclectic. They can be made from very different materials from plastic to cloth; can have hard or soft covers and pages; be very small or very large; be interactive; have dynamic elements such as in pop-up books; and even be embedded with electronics. This range of uses makes it versatile, shaping and broadening users' expectations. In this prototype, we were mainly interested in conveying the main physical characteristics of a book: its shape and size.

In this prototype, when using the Tangible VR Book, the user is initially placed in an abstract environment with only two discernible objects (Figure 4a): a screen and a stand (box). Both objects are used to present additional information beyond the content presented on the book's pages. The prototype thus uses an output coupling on object and on the environment. We also added marker-based hand tracking (Figure 4b) in order to facilitate

the task of picking the detachable parts of the book (described next) and for acting on the "slider" elements. This solution is inspired in [22] although it is clearly not optimal and is meant only for prototyping purposes. We expect that further developments in hand tracking will make it feasible in the future to have mobile-based tracking that will allow for virtual hands in the VE.
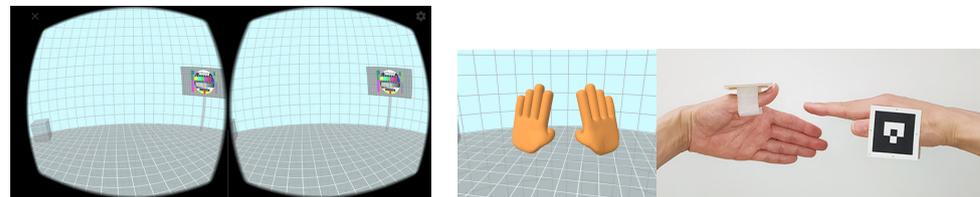
(**a**) Abstract virtual environment. (**b**) Virtual hands and markers for hand detection.

Figure 4. Abstract environment and hand detection in the Tangible VR Book #3.

The prototype of the Tangible VR Book was developed in the scope of a collaboration with the project "Digital 3D Reconstruction of the Monastery of Santa Cruz in 1834" [41]. The aim of the VR Book in this context is to allow the exploration of 360° images of the resulting digital reconstruction, as well as inspecting selected 3D models. The prototype thus uses several output modalities such as text, images, audio, video, and 3D models. Videos and 3D models are presented in the environment (the screen and stand) allowing for a larger view of the objects.

The Tangible VR Book provides three ways to browse contents (beyond the natural page flipping): through a slider activated with gestures, through a tangible slider, through tangible portals. The sliders allow in-page navigation of content, while the portals allow navigation through the 360° environments. Visual markers at the centre of each page allow the system to detect and track the book and additional markers at the bottom of the pages allow for the implementation of the content browsing mechanisms. The prototype does not have a single main I/O role, instead both input and output are equally important.

These content browsing mechanisms are important because in various situations it may not be feasible to feature a Tangible VR Book with so many pages as the number of content items to explore. To make it a more flexible solution, content browsing techniques which do not require flipping the page are required.

6.1. Gesture Slider

The gesture slider mechanism consist in performing a "sliding" gesture with the hand over the bottom part of the page, similar to touch-based ebook reading applications. This type of interaction has been explored before by Lee et al. [42]. In our implementation, the interior left page displays a table of contents and the right page displays the associated content (Figure 5). A sliding gesture left-to-right will go forward on the table of contents, while a sliding gesture right-to-left will go back. To allow displaying more content, possibly in very different formats, the browsing mechanism takes advantage of the display of information outside the book: as users browse the table of contents, videos and 3D models will be displayed on the screen and on the stand, in the environment.
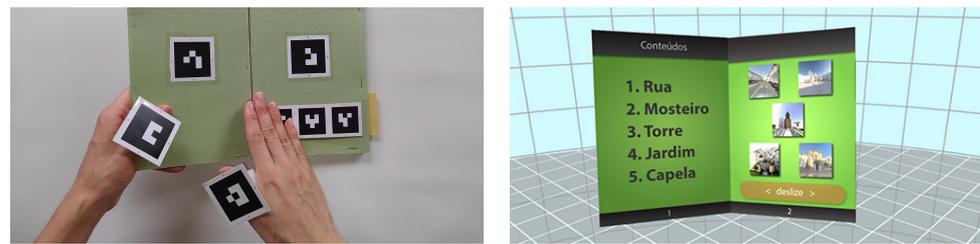
(**a**) Physical book with markers for gesture detection.

(**b**) Corresponding virtual page.

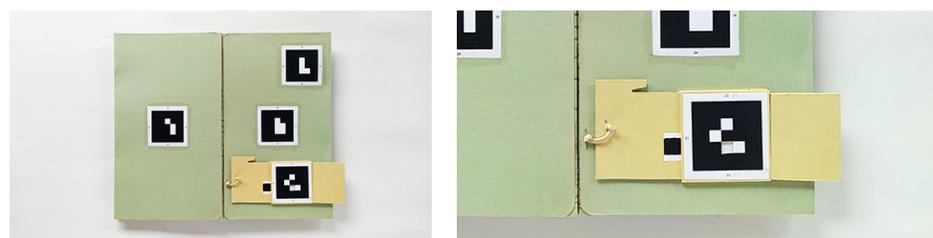**Figure 5.** Content selection through the gesture slider.

The implementation of the sliding gesture detection takes advantage of the visual markers. Three markers are placed at the bottom of the physical page and are virtually represented by a horizontal bar with the label "Slide". When users perform the sliding gesture, the system will interpret the loss of tracking of each marker in sequence as a gesture.

6.2. Tangible Slider

The tangible slider mechanism consists in physically pulling a strip of paper that slightly extends beyond the edge of the book's page (Figure 6a).

The left page shows the various content items available and, similarly to the gesture slider mechanism, the tangible slider allows users to cycle through the content by pulling and releasing the sliding strip. This will change the content depicted on the right page (Figure 6b). Zheng et al. [17] explored similar ways of creating interactable visual markers, although here we take advantage of the structure of the marker itself.

The strip is virtually represented as a small rectangular semitransparent strip at the bottom right edge of the page, along with the label "Pull" and a right facing arrow (Figure 6c). The moving strip can be made to automatically return to a resting position with the help of an elastic band. The tangible slider also provides haptic and audio cues not present in the gesture slider: when pulling the strip, users feel the resistance created by the elastic band; when the strip is released they hear the sound of the strip hitting the cardboard.

(**a**) Physical mechanism.

(**b**) Detail of the physical mechanism.
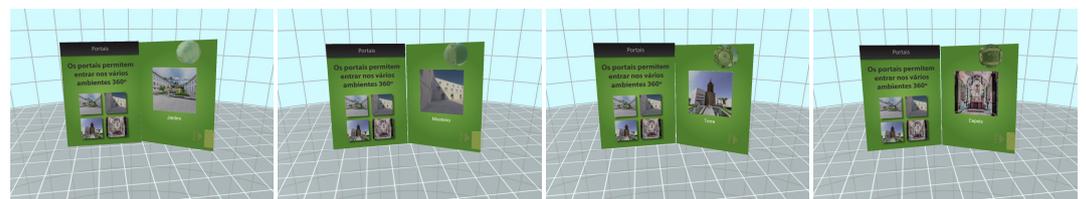
(**c**) Content items available to browse.

Figure 6. Content selection through the tangible slider.

The physical slider mechanism was specifically implemented to take advantage of the structured matrix markers [43]. Because these markers are arranged as a matrix of white and black squares, it is easy to assemble a marker, fixed to the page, with a hole that can be filled with white or black colour by an underlying moving strip (Figure 6b). This effectively

creates a dynamic object that can have two states (more states/markers could be created with holes spanning more than one cell of the matrix).

6.3. Tangible Portal

The pages with the contents for the tangible slider included also an alternative way to enter 360° environments by "picking up" a tangible portal and moving it closer to one's head as if peering inside.

The tangible for the portal is a thick block (Figure 7a) that can be removed from the top right of the page (Figure 7b). The portal is virtually represented by a semi-transparent sphere. When the user picks up the portal, a placeholder sphere is left at the top right corner of the page showing where the portal can be attached back on the page (through the use of a magnetic mechanism). After going through the portal, the VE changes automatically to display the corresponding 360° image and ambient sound.
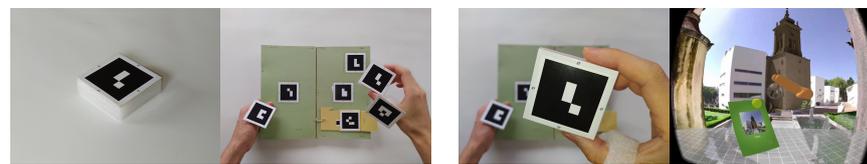
(**a**) Physical mechanism. (**b**) Going through the portal.

Figure 7. 360° navigation using the tangible portal.

6.4. Implementation

When implementing the VR Book, we aimed at creating an accessible experience that would promote generalised use instead of focusing on a single site-specific experience. We take advantage of the Computer Vision (CV) solutions for object detection and tracking often used for and associated with Augmented Reality (AR) applications. AR application have long used and improved specific CV techniques for detecting planar surfaces based on structured markers or even based on natural image features. These techniques are considered mature enough that various commercial development toolkits and platforms have emerged and are actively used even in browser platforms. Instead of implementing a typical AR application where the user holds the smartphone on her hands and points the camera to a visual marker where 3D models are then superimposed, we use the smartphone on a VR headset. The user is immersed in a VE and the detected markers are simply represented as additional 3D objects in the VE (see Figure 3).

Our implementation is based on the web-based VR framework A-Frame [44] and on the AR.js [45] component for detection of visual markers, which is itself based on a JavaScript port of ARToolkit [46]. A-Frame is web-based and supports the WebVR/WebXR specifications. It runs on smartphones, making VR experiences highly accessible and usable anywhere. The AR.js component uses the smartphone's camera to detect visual markers and calculates their position and orientation relative to the smartphone's camera (which is the same as the user's view inside the virtual world when the smartphone is placed inside the VR enclosure and put on the user's head). When a marker becomes visible or hidden on the camera, AR.js triggers an application event that is used, for example, to start/resume or to pause the video on the book's page. Multiple markers may be placed on different parts of the same physical object for creating a more robust tracking or for detecting moving parts.

The source code can be found at https://git.dei.uc.pt/jorgecardoso/Tangible-VR-Book.

7. Evaluation

Due to the COVID-19 pandemic, we opted for an evaluation procedure which would not require physically handling our prototypes. We followed an approach inspired in the vignette methodology [47] to gather feedback on our prototype. We ran two stud-

ies, through online questionnaires where users could see the prototypes being used in videos/photo-based vignettes.

In a first study, the goal was to gain additional insight on users' expectations regarding possible media content presented on the Tangible VR Book as this relates to two of the dimensions (output modality and output coupling) in our design space that are very salient to users. We also wanted to understand the general opinion regarding the usability of the prototype. In a second study, the goal was to gather feedback on the interactive elements for content navigation since they represent the least natural interactions.

We prepared online questionnaires using Google Forms and disseminated it through social networks and departmental mailing lists.

7.1. First Study

The first part of the questionnaire introduced the study and objectives and asked for demographic information (age, gender, and profession).

The second part elicited expectations regarding content and interactions: it displayed video-vignettes for the virtual representations of the various pages of a Tangible VR Book prototype and asked users several seed questions regarding what they would expect to see in the cover/back-cover of the book, as well as how they expected to interact with audio/video, 3D models, and hyperlinks inside the book. The videos in these vignettes did not show any interaction nor did they depict the user handling the book (Figure 8a). This was so as to reduce the influence on the participants since the goal was to elicit expectations.
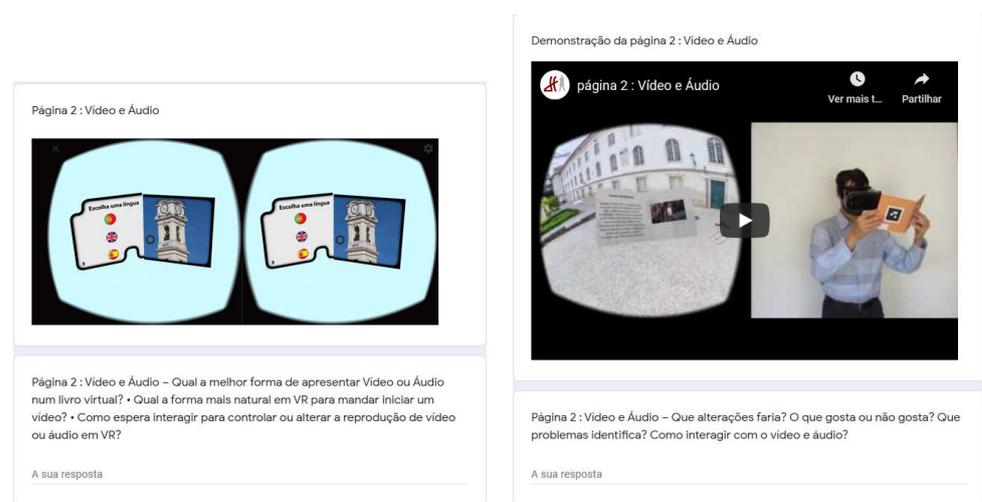
(**a**) Example vignette for the second part of the questionnaire.

(**b**) Example vignette for the third part of the questionnaire

Figure 8. Example vignettes for the second and third parts of the questionnaire.

The third part asked for feedback regarding the usability of the Tangible VR Book: it showed video-vignettes with side-by-side compositions depicting a third-person view of a user manipulating the physical Tangible VR Book and the respective first-person virtual view. Each video showed interaction with a different part of the book (interaction with video/audio, interacting with the portals, and 3D model inspection; Figure 8b). We also asked users for their general opinion regarding the prototype (disliked/not interesting, liked/interesting, liked it very much/very interesting), and additional open feedback.

The prototypes used in these vignettes were simplified versions of the prototype described in the previous section. With the intention of not drawing too much attention to these particular interactions, the prototype used for these vignettes did not show the gesture or tangible slider.

Results and Discussion

The online questionnaire was available during 18 days. Twenty one participants responded to the online questionnaire, providing over 200 comments in total.

The self-described professional activities of the respondents were "student" (5), "designer" (4), "professor" (2), "VR developer" (2), "marketing" (2), "electrical engineer" (1), "public servant" (1), "researcher" (1), "programmer" (1), "camera operator" (1), and "futurist" (1).

Ages ranged from 20 to 52 years old, with an average of 32 years old. Fourteen respondents were male and seven were female. We collected responses from several nationalities, but we did not record the country of origin of the respondent.

Expectations Regarding Content and Interactions

Ten respondents mentioned the use of video or 3D models in the cover of the book as a way to present a summary of the book's contents ("... the author, in videochroma, presenting the work."). In general, respondents' comments' pointed that dynamic content should be used in the cover as an attraction factor ("A cover with animation makes the book more appealing to the reader"). One respondent suggested that the book could trigger content outside of the book itself as a way to take advantage of the 360° space around the user ("... the opportunity of seeing in 360-hence, the reading experience of a book is no longer focused ... on the pages and becomes an immersive experience"). One respondent suggested allowing users to choose the book's layout in the cover page. This suggestion was probably inspired by current ebook readers that allow selecting the book's layout (e.g., one or two columns). However, our Tangible VR Book was not intended for high-density text display and so this feature does not currently seem relevant.

Regarding the interaction with media content (all respondents focused on video), six respondents suggested automatic video reproduction. Some participants explicitly mentioned playing automatically as the page becomes visible, other suggested playing only if the video (page) is brought closer to the user. We found it interesting that this respondent suggested this kind of interaction given that we implemented a similar interaction for the portals. It is not clear how bringing the book closer or further away to one's head could be generalised to other situations such as video playback, but it might something to explore in the future. Some respondents suggested combining automatic reproduction with manual, gesture-based commands to stop or to move the video forward or backwards. One respondent suggested an interesting shaking interaction to change the video, although it is not clear what was meant by "change": "When the page opens the video starts. To stop the video you can touch the page. To change the video you can shake the page." The shaking interaction seems interesting because it relates to a physical action that we perform with a number of physical objects, including books. The meaning of that interaction, however, is not immediate. While we do sometimes shake physical books while holding the book's covers apart in order to make as small piece of paper that we know is somewhere inside the book fall down, this is not easily applicable to other situations. Still, it is a familiar gesture that might be interesting to explore. Some participants offered suggestions inspired in their experience with touch-based devices. For example, controlling the video manually by touching the page to start and stop and pressing specific areas of the page to fast-forward or backward. Interestingly, four respondents suggested the use of voice commands. Using voice commands with a tangible interface seems contradictory at first. However, for larger virtual books where the available physical pages are not enough to display all of the virtual pages some sort of controls would be required to access the additional virtual pages. In this situation, voice input might be an alternative to consider. Three respondents suggested the use of eye-tracking, perhaps as a way to select which video would play if several were available, in combination with other mechanisms.

Regarding the interaction with hyperlinks (portals), comments were dispersed through several aspects. Three respondents suggested the use of voice commands as a way to activate the portals ("We can interact with words perhaps? Saying ENTER IN ..."). One

respondent suggested gaze-based interaction for entering a portal and another one suggested the use of specific areas in the physical page with a raised surface for haptic cue that would be pressed to enter the portal ("The page would have zones to press that would correspond to linking options, the physical page could have reliefs that allow you to feel that there is where you should press."). One respondent suggested the use of hyperlinks as a way to open up information outside VR ("Links could open a web browser close to you that shows the content.").

Regarding the use of 3D content, respondents mentioned the importance of being able to inspect and manipulate the 3D model: "I can look at a 3D model from any side, and I choose how to move the model", "Pop up models you can zoom in to, rotate or move out of the way to reveal what is behind them". One respondent suggested the ability to take the 3D model off the page ("Objects can be explored freely. I would like to touch the objects that are presented and be able to take them off the page"). Although we have not prototyped this kind of interactions, they could be implemented in smartphone-based VR with the support of other tangibles to act as grabbers or through vision-based hand detection.

Feedback Regarding Usability

Participants identified a few problems in the depicted Tangible VR Book prototype. Some were related to layout or to the visual representation of the virtual elements in the book: some pages had a lot of unused space, the video was too small, the video did not look like a video but rather like a still image, the text was too small. Respondents also noted technical problems in the detection of the markers which made pages disappear momentarily. One respondent mentioned that the background (i.e., the 360 scenario) was distracting, however others mentioned the relevance of being able to look around. One respondent considered that it might not be "very comfortable to have to hold a book in the same position to watch a long video". Two respondents considered that the sphere used to represent the portal was not intuitive: "I don't understand immediately what that sphere is for, it is not intuitive". Another respondent considered that the visual elements were not updating fast enough, which caused distraction. One respondent was also very critical regarding the value of the VR Book: "I don't see the advantage of this solution over just looking at a 2D page".

Respondents also provided several suggestions: being able to see the video in 360° instead of just on the book's page; using animated, more lively 360° backgrounds; being able to "zoom in" in the information on the book's pages by displaying the information in "a larger floating option in front of the user"; using not just 360° backgrounds but immerse the user in an explorable architectural 3D space; adding thickness to the virtual book

Some respondents also expressed their appraisal ("That sis [sic] exactly what I am looking for, to be able to see what is not there anymore") regarding the prototype and suggested further aspects for consideration. For example, one respondent highlighted the challenge of navigation within the 360° environment and wished (s)he could share the experience with other users; another suggested the use of the VR Book in a graphical adventure type of game; another suggested incorporating voice commands.

Overall Impression

When asked about their general opinion regarding the prototypes they saw while filling in the questionnaire, 20 respondents answered they liked it (10 respondents) or liked very much (10 respondents). Only one respondent disliked the Tangible VR Book prototype.

7.2. Second User Study

In the second user study, we wanted to evaluate the page content navigation techniques. In the first part of the online questionnaire, we asked for the age and gender of the participant. In the second part, participants could watch video-vignettes demonstrating the content navigation techniques, and order them by preference. The vignettes displayed a first person view of both the real and the virtual worlds and was narrated to further explain the techniques. In the third part of the questionnaire, we asked participants to

(optionally) tell us why they ordered the techniques as they did, and to provide additional comments on each technique.

In this study, we included additional page content navigation techniques beyond the two already described. In addition to the gesture slider and tangible slider, we included a touch technique where users could simply touch the bottom-left or bottom-right part of the page to slide left or right (implemented by detecting occluded markers on the bottom-left and bottom-right of the book). This served as a baseline comparison technique. Following the comments of the first study, we also included a voice technique were users could simply say "next" or "back".

Results and Discussion

We obtained a total of 50 responses, with participants' ages in the ranges [20, 29] to [50, 59] (44% of the participants were in the [20, 29] age range). Most participants (62%) were male.

Results (Figure 9) indicate the participants have a higher preference for the gesture slider technique with 24 participants selecting it as the first option, and 15 selecting it as the second option. Many participants also preferred the touch technique, with 17 participants selecting it as the first option and 22 as the second option. The tangible slider obtained the worst results, with only four participants choosing it as the first option.
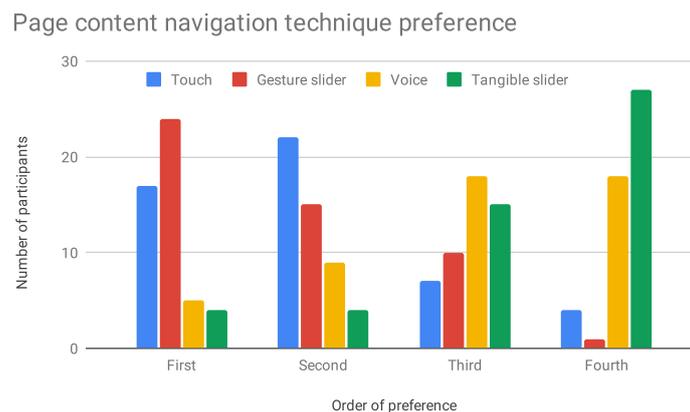

Figure 9. Page content navigation technique preference results from the second user study.

Only a few participants (five) provided explanations as to why they ordered the techniques in a given way. Regarding the gesture slider, one participant said that "for a book concept . . . [the gesture] seemed adequate", while another justified the choice with the familiarity with the gesture. We hypothesise that participants associate the sliding gesture to e-book readers and hence feel it is familiar and appropriate for a virtual book. One participant that chose the touch technique as the first option, justified it with the importance of the physical touch. Although physically touching the book was not strictly necessary, our choice of words when describing the technique implied that touching was necessary, unlike in the gesture slider technique. This participant felt that touching would be an important aspect of the experience. We were disappointed that the tangible slider was the least preferred technique. To us, this was the technique that was more in line with the idea of a tangible interface and the one that provided the richest haptic experience. However, the results seem to indicate that it is not appropriate as a page content navigation technique in this context.

Few participants provided comments specific to each interaction technique. Regarding the gesture slider, one participant suggested to use it as a way to vary the colour of some object in a more precise way. This comment seems to indicate that this participant did not fully understand how the gesture slider worked, which is comprehensible since participants only had access to a video-vignette. The gesture slider works as a discrete gesture, not as

a continuous one, which would make it difficult to use efficiently for a continuous task such as varying the colour of an object. However, this is a matter of implementation as it would be possible to create a continuous slider by calculating the distance of a marker on the slider to a fixed one. One participant suggested that the gesture slider might need some tactile element to indicate the interaction zone. Another participant commented that the gesture slider would only make sense if it mimicked the gesture of sliding our fingers through a physical page, as in turning the page of the book. These suggestions might be interesting to investigate in future versions of the Tangible VR Book.

Regarding the voice interaction, one participant commented that it could be used for other interactions but not for "changing the page". Participants also pointed out the need to inform users about the available voice commands and the importance of voice commands for users with some sort of physical disability.

Regarding the tangible slider, comments were rather vague with one participant mentioning that it displayed problems in the way it was demonstrated in the video. One participant suggested using the tangible slider to change the size of an object as in a toggle control, e.g., to display a small or a large version of a 3D model. We consider this an interesting suggesting because the tangible slider could be easily adapted to implement a control similar to a toggle button with two possible states.

8. Lessons Learned

Our early prototyping experience, together with the formative study and evaluation of the Tangible VR Book prototypes allowed us to extract lessons for future implementations of marker-based TUIs for VR.

Tracking Implementation

As pointed out by some participants in the evaluation of the Tangible VR Book, the system sometimes lost tracking of the markers, which caused flickering. Our prototypes used simple, single-marker detection and tracking algorithms which can be improved with multi-marker and extended tracking algorithms.

Going Beyond the Book Affordances Needs Careful Consideration

The portal functionality was not well understood by some participants. On the one hand, the book provides a familiar format with known affordances. On the other hand, this means that adding functionality not normally present on a book needs careful consideration and design. Additional visual cues would be necessary for users to understand the action possibilities of the portal. We believe the tangible slider suffered from a similar issue. Although it has the potential to provide a richer haptic experience, and represents a low-cost way to make a dynamic type of object, it was the least preferred option. This may be because it departs too much from how users move through content in e-book readers, for example.

Content Matters

Participants made frequent mentions to the type of content that might be featured in the cover of the Tangible VR Book. They also suggested how content on the book could be externalised in the environment. This suggests exploration of the output coupling dimension, although it is not clear if or how the physical form factor of the tangible influences this. Interaction with content is also something that provoked various kinds of suggestions. As we explained, interaction were consciously left out of the design space because they can take many forms and there are already various frameworks can be used with our design space to think about interactions.

9. Conclusions

We presented a design space of marker-based TUIs for VR built from experience gained through early prototyping and insights from a formative study. We instantiated the design space in the form of a Tangible VR Book prototype for exploring architectural

heritage. We evaluated the Tangible VR Book through and online vignette methodology inspired process to circumvent the restrictions imposed by the COVID-19 pandemic.

Our design space allows designers to consider several dimensions regarding the purpose of the marker-based TUI, facilitating the exploration of different design alternatives.

While we have mostly focused the book's content and interactions on exploration of architectural heritage, we believe the Tangible VR Book can have different uses within, and outside of, Cultural Heritage (CH).

Marker-based tangible interaction has potential as a solution for smartphone-based VR. It represents a cheap and quick way of turning physical objects into tangibles that can enhance the VR experience making it more engaging and memorable. It is also an accessible solution that can be explored in various domains and in many ways to create VR products. For example, as paper blueprints for simple objects that users can assemble; as visual markers in existing products like children's books in order to provide an alternative VR interaction modality; as marker stickers that user can stick on everyday objects, etc.

Results from the user studies on the Tangible VR Book, provided us with the perspective of users regarding acceptable types of content and what types of interactions users might expect to find. Results also provided the identification of several issues that can be correct in future versions. The results also confirm that users' attitude towards the Tangible VR Book is generally positive. This gives us confidence that it could be an interesting alternative for exploring CH contents in immersive smartphone-based VR, or, more generally, that it could be an interesting TUI for other applications.

Our user studies have obvious limitations given that participants did not actually use the tangible and could only share their opinion based on the images and videos of the system being operated by a third person. Despite this limitation, we were satisfied with the richness of the results. Participants made several suggestions regarding the use of dynamic content, specific gestures such as bringing the book close to one's head or shaking it, the use of different input modalities, the emphasis on the importance of touch, which opened up various new possibilities for the usage of the Tangible VR Book.

References

1. Ishii, H.; Ullmer, B. Tangible bits. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '97, Atlanta, GA, USA, 22–27 March 1997; Pemberton, S., Ed.; MIT Media Laboratory; ACM Press: New York, NY, USA, 1997; pp. 234–241. [CrossRef]
2. Ullmer, B.; Ishii, H. Emerging frameworks for tangible user interfaces. In *Human-Computer Interaction in the New Millennium*; Carroll, J.M., Ed.; Pearson: London, UK, 2001; Chapter 26, pp. 579–601.
3. Hinckley, K.; Pausch, R.; Goble, J.C.; Kassell, N.F. Passive real-world interface props for neurosurgical visualization. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems Celebrating Interdependence—CHI '94, New York, NY, USA, 24–28 April 1994; pp. 452–458. [CrossRef]

4. Carrozzino, M.; Bergamasco, M. Beyond virtual museums: Experiencing immersive virtual reality in real museums. *J. Cult. Herit.* **2010**, *11*, 452–458. [CrossRef]

5. Insko, B. Passive Haptics Significantly Enhances Virtual Environments. Ph.D. Thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2001.

6. Araujo, B.; Jota, R.; Perumal, V.; Yao, J.X.; Singh, K.; Wigdor, D. Snake Charmer: Physically Enabling Virtual Objects. In Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction—TEI '16, New York, NY, USA, 14–17 February 2016; pp. 218–226. [CrossRef]

7. Lee, K.M. Presence, Explicated. *Commun. Theory* **2004**, *14*, 27–50. [CrossRef]

8. Fröhlich, T.; Alexandrovsky, D.; Stabbert, T.; Döring, T.; Malaka, R. VRBox: A Virtual Reality Augmented Sandbox for Immersive Playfulness, Creativity and Exploration Thomas. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play Extended Abstracts—CHI PLAY '18, Amsterdam, The Netherlands, 15–18 October 2018; pp. 153–162. [CrossRef]

9. Podkosova, I.; Vasylevska, K.; Schoenauer, C.; Vonach, E.; Fikar, P.; Bronederk, E.; Kaufmann, H. Immersivedeck: A large-scale wireless VR system for multiple users. In Proceedings of the 2016 IEEE 9th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS), Greenville, SC, USA, 20 March 2016; pp. 1–7. [CrossRef]

10. Surale, H.B.; Gupta, A.; Hancock, M.; Vogel, D. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Scotland, UK, 4–9 May 2019; pp. 1–13. [CrossRef]

11. Johnson, M.P.; Wilson, A.; Blumberg, B.; Kline, C.; Bobick, A. Sympathetic Interfaces: Using a Plush Toy to Direct Synthetic Characters. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems the CHI is the Limit—CHI '99, Pittsburgh, PA, USA, 15–20 May 1999; pp. 152–158. [CrossRef]

12. Harley, D.; Tarun, A.P.; Germinario, D.; Mazalek, A. Tangible VR: Diegetic Tangible Objects for Virtual Reality Narratives. In Proceedings of the 2017 Conference on Designing Interactive Systems—DIS '17, New York, NY, USA, 17–19 August 2017; pp. 1253–1263. [CrossRef]

13. Sajjadi, P.; Cebolledo Gutierrez, E.O.; Trullemans, S.; De Troyer, O. Maze Commander: A Collaborative Asynchronous Game Using the Oculus Rift & the Sifteo Cubes. In Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play—CHI PLAY '14, Toronto, ON, Canada, 19–22 October 2014; pp. 227–236. [CrossRef]

14. Aguerreche, L.; Duval, T.; Lécuyer, A. Reconfigurable tangible devices for 3D virtual object manipulation by single or multiple users. In Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology—VRST '10, Lausanne, Switzerland, 15–17 September 2010; p. 227. [CrossRef]

15. Hettiarachchi, A.; Wigdor, D. Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, Jose, CA, USA, 7–12 May 2016; pp. 1957–1967. [CrossRef]

16. Jord, S.; Geiger, G.; Alonso, M.; Kaltenbrunner, M. The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In Proceedings of the TEI'07: First International Conference on Tangible and Embedded Interaction, Baton Rouge, LA, USA, 15–17 February 2007; pp. 139–146. [CrossRef]

17. Zheng, C.; Gyory, P.; Do, E.Y.L. Tangible interfaces with printed paper markers. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, New York, NY, USA, 24–28 August 2020; pp. 909–923. [CrossRef]

18. Ikeda, K.; Tsukada, K. CapacitiveMarker: Novel interaction method using visual marker integrated with conductive pattern. In Proceedings of the 6th Augmented Human International Conference, Singapore, 9–11 March 2015; Volume 11, pp. 225–226. [CrossRef]

19. Henderson, S.J.; Feiner, S. Opportunistic controls: Leveraging Natural Affordances as Tangible User Interfaces for Augmented Reality. In Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology—VRST '08, Bordeaux, France, 27–29 October 2008; p. 211. [CrossRef]

20. Paolis, L.T.D.; Aloisio, G.; Pulimeno, M. A Simulation of a Billiards Game Based on Marker Detection. In Proceedings of the 2009 Second International Conferences on Advances in Computer-Human Interactions, Cancun, Mexico, 1–7 February 2009; pp. 148–151. [CrossRef]

21. Cheng, K.Y.; Liang, R.H.; Chen, B.Y.; Liang, R.H.; Kuo, S.Y. iCon: Utilizing Everyday Objects as Additional, Auxiliary and Instant Tabletop Controllers. In Proceedings of the 28th International Conference on Human Factors in Computing Systems—CHI '10, Atlanta, GA, USA, 10–15 April 2010.

22. Lee, J.Y.; Rhee, G.W.; Seo, D.W. Hand gesture-based tangible interactions for manipulating virtual objects in a mixed reality environment. *Int. J. Adv. Manuf. Technol.* **2010**, *51*, 1069–1082. [CrossRef]

23. Billinghurst, M.; Kato, H.; Poupyrev, I. The MagicBook - Moving seamlessly between reality and virtuality. *IEEE Comput. Graph. Appl.* **2001**, *21*, 6–8. [CrossRef]

24. Ciolfi, L.; Petrelli, D.; Goldberg, R.; Dulake, N.; Willox, M.; Marshall, M.; Caparrelli, F. Exploring historical, social and natural heritage: Challenges for tangible interaction design at Sheffied General Cemetery. In Proceedings of the International Conference on Design and Digital Heritage—NODEM 2013, Stockholm, Sweden, 1–3 December 2013.

25. Lissermann, R.; Olberding, S.; Petry, B.; Mühlhäuser, M.; Steimle, J. PaperVideo: Interacting with Videos on Multiple Paper-like Displays. In Proceedings of the 20th ACM International Conference on Multimedia—MM '12, Nara, Japan, 29 October–2 November 2012; p. 129. [CrossRef]

26. Khalilbeigi, M.; Lissermann, R.; Kleine, W.; Steimle, J. FoldMe: Interacting with Double-sided Foldable Displays Mohammadreza. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction, Kingston, ON, Canada, 19–22 February 2012; pp. 33–40. [CrossRef]

27. Mazalek, A.; van den Hoven, E. Framing tangible interaction frameworks. *Artif. Intell. Eng. Des. Anal. Manuf.* **2009**, *23*, 225–235. [CrossRef]

28. Holmquist, L.E.; Redström, J.; Ljungstrand, P. Token-Based Access to Digital Information. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Gellersen, H.W., Ed.; Springer: Berlin/Heidelberg, Germany, 1999; pp. 234–245. [CrossRef]

29. Ullmer, B.; Ishii, H. Emerging frameworks for tangible user interfaces. *IBM Syst. J.* **2000**, *39*, 915–931. [CrossRef]

30. Hornecker, E.; Buur, J. Getting a grip on tangible interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing systems—CHI '06, Montreal, QC, Canada, 22–27 April 2006; p. 437. [CrossRef]

31. Simeone, A.L.; Velloso, E.; Gellersen, H. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI '15, Seoul, Korea, 18–23 April 2015; pp. 3307–3316. [CrossRef]

32. Suzuki, K.; Wakisaka, S.; Fujii, N. Substitutional Reality System: A Novel Experimental Platform for Experiencing Alternative Reality. *Sci. Rep.* **2012**, *2*, 459. [CrossRef] [PubMed]

33. Gibson, J.J. The theory of affordances. In *Perceiving, Acting and Knowing: Toward an Ecological Psychology*; Shaw, R.E., Bransford, J.D., Eds.; Wiley: New York, NY, USA, 1977.

34. Harley, D.; Chu, J.H.; Kwan, J.; Mazalek, A. Towards a Framework for Tangible Narratives. In Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction—TEI '16, Eindhoven, The Netherlands, 14–17 February 2016; pp. 62–69. [CrossRef]

35. Versteeg, M.; van den Hoven, E.; Hummels, C. Interactive Jewellery: A design exploration. In Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, Eindhoven, The Netherlands, 14–17 February 2016; pp. 44–52. [CrossRef]

36. MacLean, A.; Young, R.M.; Bellotti, V.M.; Moran, T.P. Questions, Options, and Criteria: Elements of Design Space Analysis. *Human–Comput. Interact.* **1991**, *6*, 201–250. [CrossRef]

37. Hinckley, K.; Pausch, R.; Goble, J.C.; Kassell, N.F. A survey of design issues in spatial input. In Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology—UIST '94, Marina del Rey, CA, USA, 2–4 November 1994; pp. 213–222. [CrossRef]

38. Shaer, O.; Leland, N.; Calvillo-Gamez, E.; Jacob, R. The TAC paradigm: Specifying tangible user interfaces. *Pers. Ubiquitous Comput.* **2004**, *8*. [CrossRef]

39. Benford, S.; Schnädelbach, H.; Koleva, B.; Anastasi, R.; Greenhalgh, C.; Rodden, T.; Green, J.; Ghali, A.; Pridmore, T.; Gaver, B.; et al. Expected, sensed, and desired. *ACM Trans. Comput. Hum. Interact. (TOCHI)* **2005**, *12*, 3–30. [CrossRef]

40. Norman, D.A. *The Design of Everyday Things*; Basic Books: New York, NY, USA, 2002.

41. Lobo, R. SANTA CRUZ—Reconstituição Digital do Mosteiro de Santa Cruz de Coimbra em 1834. 2018. Available online: https://santacruz.ces.uc.pt/ (accessed on 1 January 2021).

42. Lee, G.A.; Billinghurst, M.; Kim, G.J. Occlusion based interaction methods for tangible augmented reality environments. In Proceedings of the VRCAI 2004—ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry, Singapore, 16–18 June 2004; pp. 419–426. [CrossRef]

43. Garrido-Jurado, S.; Muñoz-Salinas, R.; Madrid-Cuevas, F.J.; Marín-Jiménez, M.J. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit.* **2014**, *47*, 2280–2292. [CrossRef]

44. A-Frame. 2021. Available online: https://aframe.io/ (accessed on 1 January 2021).

45. Etienne, J. AR.js. 2021. Available online: https://github.com/jeromeetienne/AR.js (accessed on 1 January 2021).

46. ARToolkit. Official ARToolkit Site. 2021. Available online: https://artoolkit.org/ (accessed on 1 January 2021).

47. Aguinis, H.; Bradley, K.J. Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organ. Res. Methods* **2014**, *17*, 351–371. [CrossRef]