



VR Book: A Tangible Interface for Smartphone-based Virtual Reality

Jorge C. S. Cardoso*

jorgecardoso@dei.uc.pt

University of Coimbra, CISUC, DEI
Coimbra, Portugal

Jorge M. Ribeiro*

jorgeribeiro@uc.pt

University of Coimbra, DEI
Coimbra, Portugal

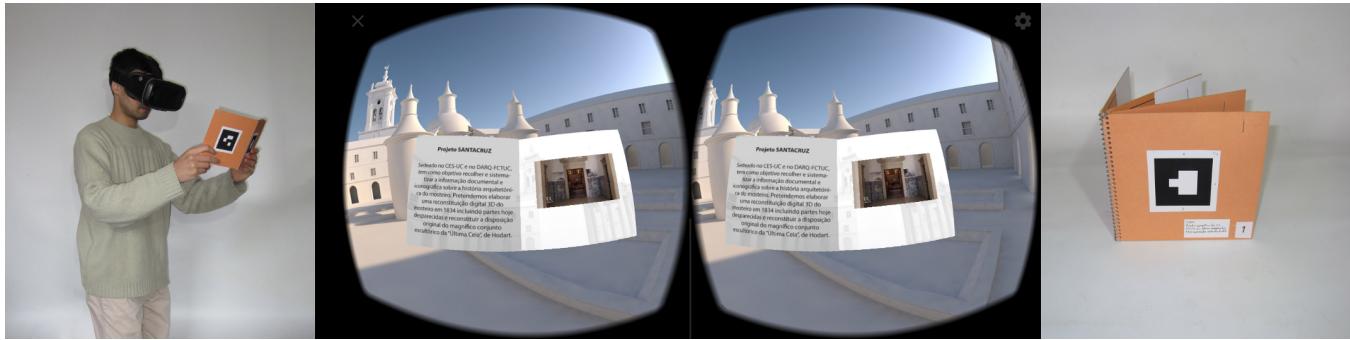


Figure 1: Overview of the Tangible VR Book.

ABSTRACT

In this work we present three prototypes of the tangible VR Book: a visual marker-based solution to tangible interaction for smartphone-based Virtual Reality (VR). Smartphone-based VR represents a low barrier to entry in VR experiences given that many people nowadays own a smartphone device and that VR headsets for these devices are affordable and quick to set up when compared to desktop-based VR. Tangible interaction in smartphone-based VR has not been much explored, in our opinion, despite the fact that it can result in easy to use and engaging experiences. In this work, we explore a marker-based solution to object tracking that allows tangibles to be created in an easy and cheap way, maintaining the overall system accessible. We describe a design space for visual marker-based tangible interaction and three prototypes of a tangible VR Book that explore different aspects of the design space. We also present user feedback on their expectations regarding the interaction with the VR Book.

CCS CONCEPTS

- Human-centered computing → Virtual reality; Smartphones; Ubiquitous and mobile computing design and evaluation methods.

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiQuitous '20, December 7–9, 2020, Darmstadt, Germany

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-8840-5/20/11...\$15.00
<https://doi.org/10.1145/3448891.3448901>

KEYWORDS

tangible user interface, tangible book, virtual reality, smartphone-based virtual reality

ACM Reference Format:

Jorge C. S. Cardoso and Jorge M. Ribeiro. 2020. VR Book: A Tangible Interface for Smartphone-based Virtual Reality. In *MobiQuitous 2020 - 17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '20), December 7–9, 2020, Darmstadt, Germany*. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3448891.3448901>

1 INTRODUCTION

Ishii & Ullmer [15] introduced the concept of Tangible User Interfaces (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.
[26]

For Virtual Reality (VR), tangible interaction represents a huge potential because tangibles naturally provide rich haptic cues which are often missing in consumer VR experiences. In VR, users often manipulate virtual objects, with obvious physical counterparts, through generic controllers. Generic controllers provide the exact same sense of weight, temperature, texture, etc., no matter what virtual object the user is manipulating. Using tangibles for VR interaction means creating a tighter association between a physical object and a virtual one: the physical object controls the virtual one (their position and orientation must match) and acts as a (haptic) representation of the virtual information. Tangible interfaces can provide more natural interactions [12], higher immersiveness [3, 14], or simply 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 – 6 Degrees of Freedom (DOF). Over the years, various alternatives have been tried, from active (instrumented with sensors), to passive objects. Many of these solutions, however, depend on some sort of infrastructure that requires a desktop computer running the VR experience, making them less portable and less adequate for walk-up-and-use situations.

We are interested in tangible interaction solutions for smartphone-based VR that do not require additional hardware instrumentation for the tangibles. In addition, we are also interested in solutions that make it easy to adapt and create tangibles. With these requirements in mind, we have started exploring the use of passive tangibles detected through marker-based computer vision algorithms such as typically employed in Augmented Reality (AR). Although there have been some applications to VR, e.g. [21], the design space for marker-based object detection for TUI VR has not been sufficiently well studied. In this work, we explore part of the design space of visual marker-based TUI VR through the implementation of a prototype Tangible VR Book. Although our main motivation for implementing the Tangible VR Book is to apply it for cultural heritage exploration in VR, we believe it can be used for several other applications.

In this paper, we briefly describe the design space for marker-based TUI for VR, the three VR Book prototypes we have developed, and their evaluation through online questionnaires.

2 RELATED WORK

Tangible interaction in VR has been realised before, with different aims and using different object tracking technologies.

2.1 Tangibles in VR

2.1.1 Active. Johnson et al. [16] instrumented 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” [9].

Sajjadi et al. [23] 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. [4] 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 [14].

Aguerreche et al. [2], 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 [11] for example, a Kinect sensor is used 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 Visual markers in VR / AR

In this work, our approach is to take advantage of the mature body of work on computer vision algorithms for detection of visual markers that are typically employed for AR applications, and apply it to detect physical objects using smartphone-based VR.

Visual markers have been used extensively for AR applications but there are not many examples of visual markers used in immersive VR. Although not strictly for VR, the following examples demonstrate the possible uses of visual markers for detecting objects and provide good examples of how they could be used in smartphone-based VR.

Henderson & Steven [10] 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” [10]. 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 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. [6] in their iCon system, stuck fiducial markers on everyday objects to convert them into input controllers for various applications.

Lee et al. [17] 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 one of our prototypes for the VR Book, we experimented with visual markers on the user's hands as a way to provide visual feedback about the position of the hands.

3 VISUAL MARKER-BASED TANGIBLE INTERACTION IN VIRTUAL REALITY

3.1 General approach and implementation

When implementing the Tangible VR Book, we aimed at creating an accessible experience what would promote generalized 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 1).

Our implementation is based on the web-based VR framework A-Frame [1] and on the AR.js [7] component for detection of visual markers, which is itself based on a JavaScript port of ARToolkit [5]. 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. To trigger a jump through a portal based on the distance to the user's head, for example, we continually calculate the distance between the virtual camera and the 3D object that represents the portal which in turn is attached to the tangible. When this distance reaches a minimum threshold, the application loads a new scene.

Multiple markers may be placed on different parts of the same physical object for creating a more robust tracking or for detecting moving parts. Figure 2 shows an initial prototype for an architectural model of a tower that users could "open up" to inspect its interior. Several markers were placed on the outside and inside of the cardboard box to allow the object to be inspected from all sides.

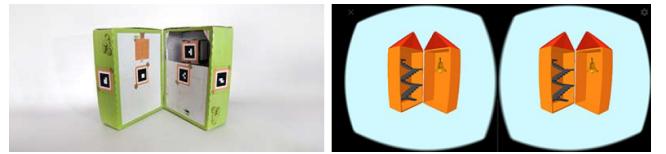


Figure 2: Tangible tower prototype. Left: physical cardboard object with several markers; right: user's perspective of the virtual representation.

In the interior, a marker was also placed on a suspended piece of cardboard (with an additional metal piece for weight) that simulates the tower's bell. The natural physical movement of the cardboard box will make the virtual bell swing.

3.2 Design space

As part of this work, we characterized a design space for marker-based tangibles for VR. A detailed account of the process through which we arrived at this design space is outside the scope of this paper, but it included a literature review, brainstorming sessions, and interviews with professionals from different areas. The result is a set of 8 dimensions that provide different design options (Table 1). We use the term "dimensions" as synonym to the "questions" of MacLean et al. [18] in their Design Space Analysis method.

Table 1: Design space for marker-based tangibles for VR: dimensions and options

Dimension	Options
Overall Fidelity	Proxy, Replica
Main physical characteristic	Smell, Temperature, Texture, Weight, Size, Rigidity, Shape
Main I/O role	Output (Display), Input (Manipulation, Controller)
Output modality	3D Model, Video, Audio, Image, Text
Output coupling	Environment, Coupled, On object
Type of object	Static/Dynamic, Passive/Active, Non-Reconfigurable/Reconfigurable, Immovable/Movable
System's Configuration	Associative, Constructive, Relational, Spatial
Informational role	Tools, Tokens, Containers

3.2.1 Overall fidelity. The *overall fidelity* dimension captures part of the layered model of substitutional environments by Simeone et al. [24]. We simplify the layered model however, by considering only two options: *replica* – a reproduction of an existing physical

object, and *proxy* – includes the remaining layers of substitutional environments: aesthetic, addition/subtraction, function, category.

3.2.2 Main physical characteristic. The *main physical characteristic* dimension captures the most relevant physical features when designing the tangible object for a particular VR experience: *shape, rigidity, size, weight, texture, temperature, smell, sound*.

3.2.3 Main I/O role. Although tangibles are always used both as input to the system 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.

3.2.4 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). Most often the tangible is represented through a *3D model* in the VE, but other options are possible: *text, image, audio, video*.

3.2.5 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.

3.2.6 Type of object. The *type of object* dimension combines several binary properties that the physical object can exhibit: *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 a tangible for a book would be considered reconfigurable if it had pages that 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.

3.2.7 System configuration. The *system configuration* dimension is directly taken from and equivalent to the categories defined by Ullmer and Ishii [25]: *spatial, relational, constructive, and associative*.

3.2.8 Informational role. The *informational role* dimension is equivalent to Holmquist et al.'s [13] classes of physical objects that represent digital information: *containers, tokens, and tools*.

4 TANGIBLE VR BOOK

The Tangible VR Book (Figure 1) is composed by a simplified physical book (a few pages only) with thick pages with visual AR markers on each page. This roughly keeps the same affordances [19] of a normal book: it can be picked up for inspecting the cover and back cover, pages can be turned, it can be brought closer or farther away from one's eyes. These affordances are the same for the virtual representation of the VR book.

4.1 VR Book #1

The first 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”. 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 (Figure 3).

The cover contains the title of the project and an image of a render of the digital reconstruction of the monastery (Figure 3a). The interior pages include different types of media beyond static images and text. For example, in the second interior page, a video is automatically played when the page is opened, and paused when the page is closed (Figure 3b). The fourth (right) interior page depicts a portal to the location described in the left page. The portal is represented as a textured sphere and it can be triggered by bringing the book closer to one's eyes (Figure 3c). Once triggered, the user will be instantly teleported to that location (a 360° scene). Although this prototype depicts only one portal, there could be several, in different interior pages providing a walkthrough through the main reconstructed areas, highlighting relevant aspects in the text (left page), displaying static rendered images, and providing a portal to the 360° view of that area. This provides users with a natural interaction to quickly go through the various 360° renderings without using artificial VR controllers. Finally, 3D content can also be displayed, similarly to a physical popup book, with the added possibility of including 3D animated models (Figure 3d).

4.2 VR Book #2

The second VR Book prototype explored alternatives to the contents depicted in the various pages and alternative interaction techniques, essentially based on gaze-cursor (reticle) interaction. For this prototype, we created a book to explore the campus of the University of Coimbra.

The cover page (Figure 4a (left)), shows a static image of the University Palace with title text in a different plane, floating above the image. It also includes an animation of the “730 years” caption at the bottom. The interior pages 1 and 2 (Figure 4a (right)) show an iconic building of the University – its Tower – as an interactive 3D model. Gazing at the tower will trigger the sound of the tower's bell. In the interior pages 3 and 4 (Figure 4b), the book shows a map of the campus of the University and several popout pins identify the locations of the various buildings. Gazing on one of those pins will display a photo of the respective building overlaid on the page. Pages 5 and 6 allow users to hear and see a video about the University, narrated in one of three languages. Users can select the language by gazing at the respective flag icon on the left page (Figure 4c). Finally, the back-cover of the book shows an animated



a) Cover and back cover.



b) Text and video.



c) Portal.



d) 3D Object.

Figure 3: Details of the virtual pages of the tangible VR Book #1 prototype for the Santa Cruz project.

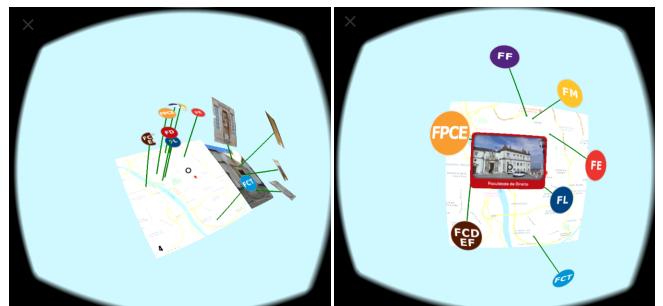
3D model of a closed window that will display a view of the Palace yard when opened (Figure 4d).

4.3 VR Book #3

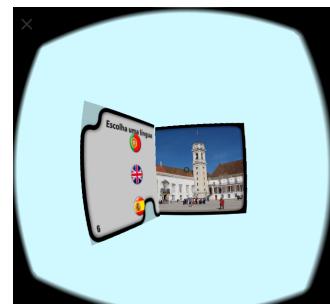
The third VR Book prototype represents an attempt to push the limits of the design incorporating options from the design space that



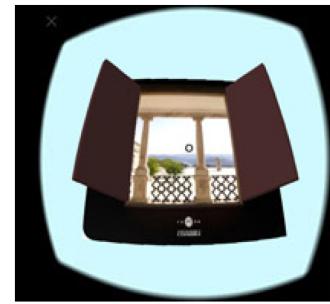
a) Cover (left): Depicts the University Palace with standing out text and animation. Interior pages 1 and 2 (right) Depicts the University's Tower with a 3D model; Gazing at the tower will trigger the bell's sound.



b) Interior pages 3 and 4; A map of the University with pins on the locations of the main buildings; Gazing at the pins will display the photo of the building.



c) Interior pages 5 and 6; A video about the University narrated in the language chosen by the user.



d) Back cover: 3D objects animated when gazed.

Figure 4: Details of the virtual pages of the tangible VR Book prototype about the University of Coimbra.

were not addressed by the two previous prototypes: display of information associated with the book on the environment (instead of on the book itself); use of a system of tangibles, including dynamic objects, together with the book to allow browsing content beyond the physical limits of the number of pages, and; virtual representation of the user's hands. It is important to note that this prototype was developed after the user feedback described in the Evaluation section and some of the implemented features resulted directly from that feedback. This prototype was not evaluated though.

In this prototype, the user is initially placed in an abstract environment with only two discernible objects (Figure 5): a screen and a stand (box). Both objects are used to present additional information beyond the content presented on the book's pages.

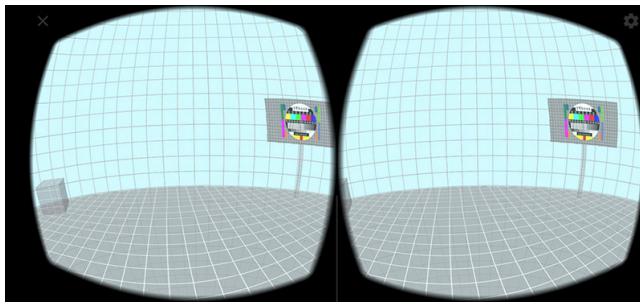


Figure 5: Abstract virtual environment for VR Book #3.

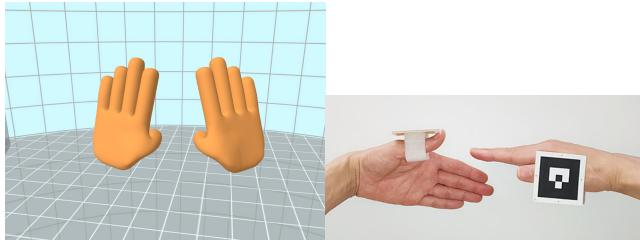


Figure 6: Virtual hands and markers for hand detection in VR Book #3.

For this prototype, we also added marker-based hand tracking (Figure 6) 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 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.

It should be stressed that this prototype was developed purposefully to push the limits of the previous prototypes, in what concerns marker-based interaction, and to allow reflection on the kinds of features that might make sense in such a tangible. This prototype was not meant as a direct basis for a final product.

This VR Book prototype provides three ways to browse contents (beyond the natural page flipping): through a slider activated with gestures, through a tangible slider, through tangible portals. Like in the previous prototypes, visual markers at the center of each

page allow the system to detect and track the book. However, in this prototype, additional markers allow for the implementation of the content browsing mechanisms.

4.3.1 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 (although in the VR Book it is not strictly mandatory to physically touch the page). In our implementation, the interior left page displays a table of contents and the right page displays the associated content (Figure 7). 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.



Figure 7: 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.

4.3.2 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 8a).

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 depicted on the right page (Figure 8b). Unlike the gesture slider, this implementation of the tangible slider allows browsing in only the forward direction. 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 8c).

The physical slider mechanism was specifically implemented to take advantage of the structured matrix markers [22]. 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 8b). 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). The moving strip can be made to automatically return to a resting position with the help of an elastic band.

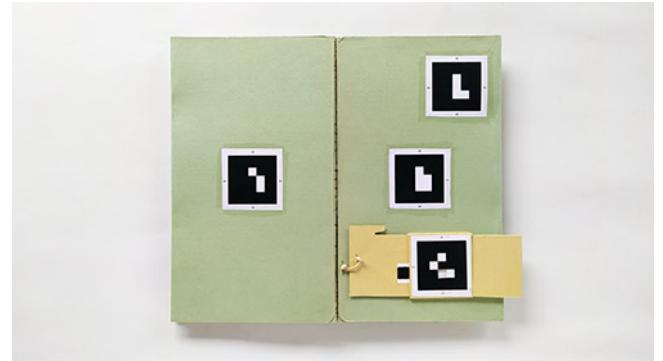
4.3.3 Tangible portal. The pages with the contents for the tangible slider included also an alternative way to enter 360° environments by “picking up” a tangible portal and moving it closer to one’s head as if peering inside.

The tangible for the portal is a thick block (Figure 9a) that can be removed from the top right of the page (Figure 9b). 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 magnet mechanism). After going through the portal, the VE changes automatically to display the corresponding 360° image and ambient sound.

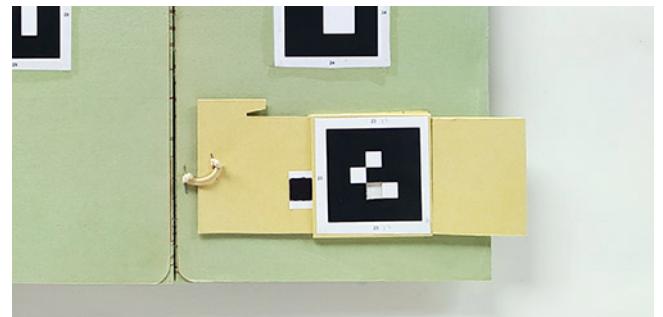
5 EVALUATION

Evaluation of the VR Book prototypes was impacted by the 2020 Covid-19 pandemic and adapted to the existing restrictions regarding social distancing and partial shutdown of Universities. While we had originally planned to perform usability testing with users in our lab, this became unfeasible given the mobility and access restrictions that were in place. We also considered remote usability testing, but this would require us to either send materials (mobile VR headset and cardboard VR Book prototype), or recruit only users that owned a mobile VR headset and ask them to build the book prototype according to our instructions. Given the safety concerns of the first option, and required effort for the second, we instead opted for an alternative way of gathering user feedback. We collected feedback through online questionnaires where users could see the prototypes being used in videos, photos and animated GIFs.

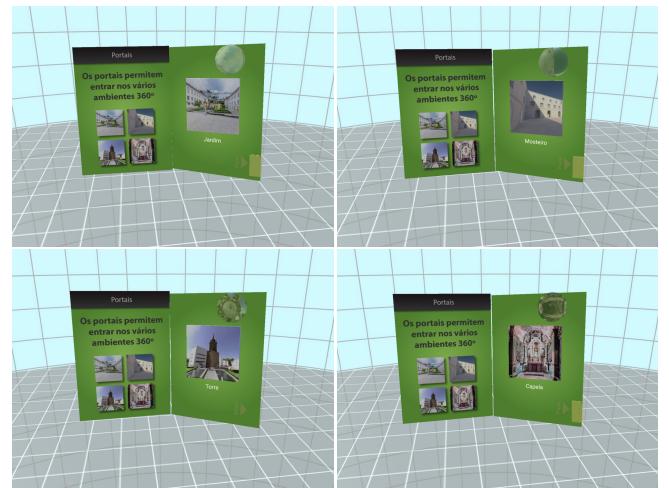
We prepared an online questionnaire using Google Forms and disseminated it through social networks, departmental mailing lists, and personal contacts. We collected feedback on the user expectations regarding possible media content presented on the VR Book, feedback on the VR Book #1, and the general opinion regarding the prototypes.



a) Physical mechanism.



b) Detail of the physical mechanism.



c) Content items available to browse.

Figure 8: Content selection through the tangible slider.

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

The second part elicited expectations regarding content and interactions: it displayed animated GIFs for the virtual representations of the various pages of VR Book #2 and asked users:

- Which multimedia content to you think could be used on the book’s cover?

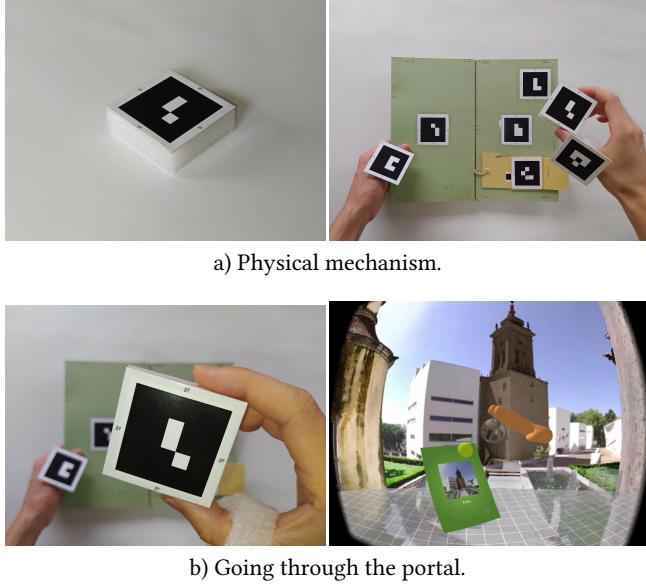


Figure 9: 360° navigation using the tangible portal.

- What is the best way to present video or audio in a virtual book?
- What is the most natural way to start a video in VR?
- How would you expect to be able to control the video or audio reproduction?
- In what way do you expect to be able to use hyperlinks in a virtual environment?
- Given the various faculties represented on the page, how do you expect to be able to interact to “enter” that option?
- To what other kind of content could hyperlinks be of use in VR?
- What do you consider the advantages of interaction in virtual reality with the object to manipulate 3D models?
- What kind of 3D models would you like to manipulate/explore with this form of interaction in VR?
- What ways of interaction would you like to have with a 3D model in VR?
- On the back cover of the book, how do you imagine being able to interact to choose a new book to explore with the same tangible support?
- What kind of multimedia and interaction content would be best to end the book with?

These questions were grouped and presented next to animated GIFs of each page set in the book. However, the purpose of the various animated GIFs were mostly to provide users with a “seed” to think about the kinds of content and interactions and not to gather specific feedback about the content depicted in the GIFs.

The third part asked for feedback regarding VR Book #1: it showed three videos with side-by-side compositions depicting a third-person view of a user manipulating the physical VR Book #1 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). After each video, users were asked to answer:

- Video and Audio – What changes would you make? What do you like or dislike? What problems do you identify? How to interact with video and audio?
- Portals – What do you like or dislike? How to interact with links? What problems do you identify? What multimedia content or spaces would you use in this context?
- 3D models – What do you like or dislike? What 3D models would be interesting to explore? How to interact with the elements? What problems do you identify? What potential do you see in this method?

Finally, the fourth part of the questionnaire gather the general opinion regarding the prototypes (Disliked / not interesting, Liked / interesting, I liked it very much / very interesting), and open feedback regarding multimedia content, interaction, tangible properties of the physical object, or any other comments participants would like to make.

5.1 Results and Discussion

The online questionnaire was available during 18 days during April and May 2020. Twenty one participants responded to the online questionnaire, providing over 200 comments in total.

5.1.1 Demographics. 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.

5.1.2 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.”¹). 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”²). 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”³). We explored this suggestion in VR Book #3 by displaying video and 3D models in the environment instead of the book itself. However, further evaluation would be required to assess if this feature would be valued by users in any particular circumstance. 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 VR Book was not

¹Translated from “... autor, em videochroma, a apresentar a obra.”

²Translated from “Uma capa com animação torna o livro mais apelativo ao leitor.”

³Translated from “... a oportunidade de vermos em 360 - logo, a experiência de leitura de um livro deixa de estar concentrada ... nas páginas e é uma experiência imersiva.”

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 generalized 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. While we did not implement gestures for video control, in VR Book #3 we implemented gestures for loading new content on the same page. 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."⁴. 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 a small piece of paper that we know is somewhere inside the book fall down, this is not easily applicable to other situations. 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 ..."⁵). 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."⁶). One respondent suggested the use of hyperlinks as a way to open up

⁴Translated from "Quando abrir a página inicia o vídeo. Para parar o vídeo podes tocar na página. Para alterar o vídeo podes abanar a página."

⁵Translated from "Podemos interagir com palavras tal vez ? Dizendo ENTRAR EM"

⁶Translated from "A página teria zonas para pressionar que correspondem a possíveis opções de ligações, a página física pode ter relevos que fazem sentir que é nessa posição que deverá pressionar."

information outside VR ("Links could open a web browser close to you that shows the content").

Regarding the use of 3D content in the VR Book, 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"⁷, "Pop up models you can zoom in to, rotate or move out of the way to reveal what is behind them". One respondent mentioned 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").

5.1.3 Feedback regarding VR Book #1. Regarding this part of the questionnaire, most respondents did not directly answer most of the questions (which were all optional), but they identified several problems in the VR Book #1. Many of the issues problems 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 the 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.

5.1.4 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 VR Book prototypes.

⁷Translated from "Posso ver um modelo 3d de qualquer lado, e eu escolho como mover o modelo."

6 CONCLUSIONS

We have started to explore the design space of marker-based tangible interaction for smartphone-based VR. We have characterized the design space along eight dimensions with several options that can help choose design alternatives and we have implemented three variations on a tangible VR Book that explores different aspects of the design space. We have gathered user feedback through an online questionnaire regarding the expectations users have about the interaction with content on the VR Book. User feedback allowed us to consider alternative interactive features that we prototyped in a third VR Book.

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 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.

Our evaluation of the VR Book has obvious limitations. We evaluated the VR Book prototypes through online questionnaires in which respondents experienced the prototypes only through a textual description, still images, and videos. No user experienced the prototypes first-hand. On the one hand, this has certainly resulted in a few misunderstandings about the behaviour and interaction with the system. On the other, we were unable to ask respondents for clarifications regarding their comments. In hind-sight, we could have performed a few synchronous remote sessions, which would have allowed us to interview participants afterwards.

Our initial motivation for marker-based tangibles for VR was for the exploration of digital reconstructions in cultural heritage in which the efficiency of the interaction is not the most important factor driving the design. However, in future work we will compare marker-based tangible interaction with alternative forms of interaction and assess different quality factors in the VR experience.

ACKNOWLEDGMENTS

This work was financed by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operacional Programme for Competitiveness and Internationalisation (POCI), and by Portuguese funds through FCT - Foundation for Science and Technology, I.P., in the framework of the project 30704 (Reference: POCI-01-0145-FEDER-030704)

This work is funded by national funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020.

REFERENCES

- [1] A-Frame. 2019. A-Frame. <https://aframe.io/>
- [2] Laurent Aguerreche, Thierry Duval, and Anatole Lécuyer. 2010. 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*. ACM Press, New York, New York, USA, 227. <https://doi.org/10.1145/1889863.1889913>
- [3] Bruno Araujo, Ricardo Jota, Varun Perumal, Jia Xian Yao, Karan Singh, and Daniel Wigdor. 2016. Snake Charmer: Physically Enabling Virtual Objects. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '16*. ACM Press, New York, New York, USA, 218–226. <https://doi.org/10.1145/2839462.2839484>
- [4] Bruno Araujo, Ricardo Jota, Varun Perumal, Jia Xian Yao, Karan Singh, and Daniel Wigdor. 2016. Snake Charmer: Physically Enabling Virtual Objects. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '16*. ACM Press, New York, New York, USA, 218–226. <https://doi.org/10.1145/2839462.2839484>
- [5] ARToolkit. 2018. Official ARToolkit site. <https://artoolkit.org/>. <https://artoolkit.org/>
- [6] Kai-Yin Cheng, Rong-Hao Liang, Bing-Yu Chen, Rung-Huei Liang, and Sy-Yen Kuo. 2010. 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*. ACM Press, New York, New York, USA. <http://www.violet.net/>
- [7] Jerome Etienne. 2020. AR.js. <https://github.com/jeromeetienne/AR.js>
- [8] Thomas Fröhlich, Dmitry Alexandrovsky, Timo Stabbert, Tanja Döring, and Rainer Malaka. 2018. VRBox: A Virtual Reality Augmented Sandbox for Immersive Playfulness, Creativity and Exploration Thomas. In *The Annual Symposium on Computer-Human Interaction in Play Extended Abstracts - CHI PLAY '18*. ACM Press, New York, New York, USA, 153–162. <https://doi.org/10.1145/3242671.3242697>
- [9] Daniel Harley, Aneesh P. Tarun, Daniel Germinario, and Ali Mazalek. 2017. Tangible VR: Diegetic Tangible Objects for Virtual Reality Narratives. In *Proceedings of the 2017 Conference on Designing Interactive Systems - DIS '17*. ACM Press, New York, New York, USA, 1253–1263. <https://doi.org/10.1145/3064663.3064680>
- [10] Steven J. Henderson and Steven Feiner. 2008. 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*. ACM Press, New York, New York, USA, 211. <https://doi.org/10.1145/1450579.1450625>
- [11] Anuruddha Hettiarachchi and Daniel Wigdor. 2016. 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*. ACM Press, New York, New York, USA, 1957–1967. <https://doi.org/10.1145/2858036.2858134>
- [12] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94*. ACM Press, New York, New York, USA, 452–458. <https://doi.org/10.1145/191666.191821>
- [13] Lars Erik Holmquist, Johan Redström, and Peter Ljungstrand. 1999. Token-Based Access to Digital Information. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Hans-W. Gellersen (Ed.). Springer, 234–245. https://doi.org/10.1007/3-540-48157-5_22
- [14] Brent Insko. 2001. *Passive Haptics Significantly Enhances Virtual Environments*. Ph.D. Dissertation. University of North Carolina at Chapel Hill.
- [15] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '97 (CHI '97, March)*, Steven Pemberton (Ed.), MIT Media Laboratory, ACM Press, New York, New York, USA, 234–241. <https://doi.org/10.1145/258549.258715>
- [16] Michael Patrick Johnson, Andrew Wilson, Bruce Blumberg, Christopher Kline, and Aaron Bobick. 1999. 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*. ACM Press, New York, New York, USA, 152–158. <https://doi.org/10.1145/302979.303028>
- [17] Jae Yeol Lee, Gue Won Rhee, and Dong Woo Seo. 2010. Hand gesture-based tangible interactions for manipulating virtual objects in a mixed reality environment. *The International Journal of Advanced Manufacturing Technology* 51, 9–12 (12 2010), 1069–1082. <https://doi.org/10.1007/s00170-010-2671-x>
- [18] Allan MacLean, Richard M. Young, Victoria M.E. Bellotti, and Thomas P. Moran. 1991. Questions, Options, and Criteria: Elements of Design Space Analysis. *Human–Computer Interaction* 6, 3–4 (1991), 201–250. <https://doi.org/10.1080/07370024.1991.9667168>
- [19] Donald A Norman. 2002. *The Design of Everyday Things*. Basic Books.
- [20] Lucio T. De Paolis, Giovanni Aloisio, and Marco Pulimeno. 2009. A Simulation of a Billiards Game Based on Marker Detection. In *2009 Second International Conferences on Advances in Computer-Human Interactions*. IEEE, 148–151. <https://doi.org/10.1109/ACHI.2009.19>
- [21] Iana Podkosova, Khrystyna Vasylevska, Christian Schoenauer, Emanuel Vonach, Peter Fikar, Elisabeth Bronederk, and Hannes Kaufmann. 2016. Immersedeck: a large-scale wireless VR system for multiple users. In *2016 IEEE 9th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*. IEEE, 1–7. <https://doi.org/10.1109/SEARIS.2016.7551581>

- [22] Jun Rekimoto. 1998. Matrix: a realtime object identification and registration method for augmented reality. In *Proceedings. 3rd Asia Pacific Computer Human Interaction*. IEEE Comput. Soc, Shonan Village Center, Japan, 63–68. <https://doi.org/10.1109/APCHI.1998.704151>
- [23] Pejman Sajjadi, Edgar Omar Cebollo Gutierrez, Sandra Trullemans, and Olga De Troyer. 2014. 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*. ACM Press, New York, New York, USA, 227–236. <https://doi.org/10.1145/2658537.2658690>
- [24] Adalberto L. Simeone, Eduardo Veloso, and Hans Gellersen. 2015. Substitutional Reality. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press, New York, New York, USA, 3307–3316. <https://doi.org/10.1145/2702123.2702389>
- [25] B. Ullmer and H. Ishii. 2000. Emerging frameworks for tangible user interfaces. *IBM Systems Journal* 39, 3.4 (7 2000), 915–931. <https://doi.org/10.1147/sj.393.0915>
- [26] Brygg Ullmer and Hiroshi Ishii. 2001. Emerging frameworks for tangible user interfaces. In *Human-computer interaction in the new millennium*, John M. Carroll (Ed.). Pearson, Chapter 26, 579–601.