A Holistic User-Centered Approach to Immersive Digital Cultural Heritage Installations: Case Vrouw Maria

MARKKU REUNANEN, LILY DÍAZ, and TOMMI HORTTANA, Aalto University School of Arts, Design and Architecture

This article deals with the design and implementation of an immersive installation where users could gesturally navigate around the wreck of Vrouw Maria, a Dutch merchant ship that sank near the Finnish coast in 1771 and was rediscovered in 1999. The installation was built for the Maritime Museum in Kotka, Finland, and is part of the preservation efforts of the wreck, which still remains underwater. In addition to the cultural heritage aspect, the project was an experiment in holistic user-centered design, where several design methods, such as scenarios, role playing and informance, storyboards, and prototyping, were employed throughout the process in order to envision the final product as well as assess their utility in the scope of immersive installations. The approach we have taken and documented here can be used as a starting point for similar projects where archaeological sites are reconstructed virtually and presented, for example, in a museum setting.

Categories and Subject Descriptors: H.5.1 Information Systems [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Artificial, augmented, and virtual realities*; H.5.2 Information Systems [**Information Interfaces and Presentation**]: User Interfaces—*Interaction styles*; H.5.2 Information Systems [**Information Interfaces and Presentation**]: User Interfaces—*User-centered design*

General Terms: Human-Centered Computing

Additional Key Words and Phrases: Bodily interaction, cultural heritage, installation

ACM Reference Format:

Markku Reunanen, Lily Díaz, and Tommi Horttana. 2015. A holistic user-centered approach to immersive digital cultural heritage installations: Case Vrouw Maria. ACM J. Comput. Cult. Herit. 7, 4, Article 24 (February 2015), 16 pages.
DOI: http://dx.doi.org/10.1145/2637485

1. INTRODUCTION

Vrouw Maria was a Dutch 18th-century merchant ship that was on its way from Amsterdam to Saint Petersburg when it hit rocks in a storm and sank near Nauvo, Finland (part of the Swedish empire at the time) in 1771. The crew managed to escape and salvage some of the cargo, but most of it sank with the ship. The 26-meter-long wreck was rediscovered in 1999 and has been extensively studied since then. As of now, the hull of *Vrouw Maria* is relatively well preserved, owing to the lack of wood-eating shipworms in the Gulf of Finland. Only a few items have been retrieved by scuba divers for research purposes, so most of the cargo still remains in the wreck, contributing to its “treasure ship” myth. For

Authors’ address: M. Reunanen, L. Díaz, and T. Horttana, Aalto University School of Arts, Design and Architecture, Department of Media, Miestentie 3, 02150 Espoo, Finland; email: {markku.reunanen, lily.diaz, tommi.horttana}@aalto.fi.

DOI: http://dx.doi.org/10.1145/2637485

a more complete overview of the history of *Vrouw Maria* and the ongoing preservation efforts, see *Lost at Sea, Rediscovered* [Ehanti et al. 2012].

In addition to being protected by Finnish legislation concerning archaeological sites, the area is a natural reserve, which makes it impossible for the public to visit the site. Furthermore, the wreck lies at a depth of 41 meters where the visibility is constantly low, so only experienced divers can reach it safely. Thus, a virtual tour is the only option to visit the ship before it is lifted—which might never happen—and painstakingly conserved for display at a museum. In this article, we discuss the design and implementation of an immersive installation that allows museumgoers to interactively experience the site, navigate its surroundings, and find out about its history by visiting info spots located at various points of interest. The installation was built as part of the *Vrouw Maria Underwater* project that ran from 2009 to 2012 and finally deployed at the Maritime Museum in Kotka in 2012. The main features are as follows (see Section 5 for further details):

—Real-scale graphics based on images and measurements obtained from the site
—Immersive stereoscopic display consisting of a back-projected screen (280 × 175 cm) and polarized goggles
—Gesture-based interaction
—Immersive soundscape that changes according to time and depth
—15 info spots offering further details on the site

Technology drive and a desire to push beyond the frontiers of the mainstream in new media, particularly with respect to the collaborative development of a gesture-based interface, were some of the factors behind the development of the *Vrouw Maria* installation. Creating new venues for the representation of the artistic, historical, and scientific narratives embedded within the heritage site was another. Finally, providing access to the previously hidden and inaccessible phenomena and artifacts present at the site was also a prime motivation.

The approach we took in the case of the Vrouw Maria installation was to employ well-known user-centered design methods holistically throughout the whole process, starting from the early concepting phase to the final user test conducted with an almost complete installation. In addition to the aim of building an engaging, rich user experience, the project serves as an example case where the usefulness of different design and prototyping methods was evaluated in the context of immersive digital cultural heritage installations—maritime archaeology in particular. Several domain experts, such as archaeologists and divers, participated in the project along the way, which let us observe their role and contribution at different stages of design and implementation.

2. DIGITAL CULTURAL HERITAGE AND IMMERSIVE INSTALLATIONS

Digital cultural heritage (DCH) describes tangible and intangible cultural heritage entities that have been either created in or transformed to a digital media format. The term also refers to novel applications that make use of digital media and information processing tools to allow for new ways of displaying, manipulating, working with, and storing cultural heritage artifacts. Throughout its brief history, the research agenda and knowledge objectives in digital cultural heritage have been influenced by developments in the field of information technology and computer science. Beginning in the early 1980s, cultural heritage and contents from memory institutions—including some of the world's most noteworthy—have served both as inspiration and as a testing ground of digital tool development. Nowadays, digital cultural heritage is an interdisciplinary area of knowledge involving the research, study, and analysis of the digital representations of heritage and how these are created and used in the preservation and dissemination of culture.

Virtual reality (VR) technology is concerned with the creation of immersive, artificially produced reality experiences. In many cases, digital cultural heritage applications have sought to use VR technology to provide visitors with substitute yet fulfilling experiences of visiting an actual heritage site. One of the earliest examples is the *Palenque Project*, clearly inspired by MIT's *Aspen Movie* project and developed in 1985 by the Bank Street College and David Sarnoff Research Center, as part of the demonstration of new interactive digital video technology. The application combined panorama photography with motion video and sound and allowed for a virtual tour of the famed site [Wilson 1988]. *Palenque's* panorama-based approach has continued to be developed, one of the latest examples being *Place-Hampi*, a groundbreaking new media installation. The installation combines stereographic photography in a 360-degree panorama screen with motion video and a spatial sound system to engage the visitors in a multisensory journey of discovery of the World Heritage site of the Ancient Hindu Kingdom of Vijayanagara in South India [Place-Hampi 2008].

Bartley and Hancock [2008] discuss the rich possibilities that reconstructions could offer to memory institutions in the form of virtual tourism. They describe the experience as an *embodied spatiality*, a sense of place, that allows for direct participation with the subject matter presented. As can be witnessed in the works showcased by the European Union–funded *Virtual Museum Transnational Network*, these types of installations and environments are becoming a new form of computer-mediated communications that have the potential to offer museum visitors novel types of multimodal experiences of cultural heritage [V-Must 2011].

Some DCH installations have blended real environments with virtual content by using Augmented Reality (AR) technology. Conceptually, AR resides between pure virtuality and pure reality on the mixed reality continuum, which was first presented by Milgram et al. [1994]. Such installations tend to be highly site specific, since they rely on the physical properties of a particular archaeological site. *ARCHEOGUIDE*, first deployed at Olympia, Greece, is one of the earliest examples of applying AR to a DCH context [Vlahakis et al. 2001]. *The Last Supper Interactive* project, based on Leonardo da Vinci's famous mural, demonstrates the malleability of digital media, as it is available both as a standalone VR installation and as a site-specific mobile AR application [F.A.B.R.I.CATORS 2011].

The main context of this study is DCH in maritime archaeology, often characterized by hard-to-access underwater sites. Building miniatures and 3D models of shipwrecks is a common way of improving the accessibility: for instance, already before our virtual reconstruction, both professionals and hobbyists had built several miniature models of *Vrouw Maria*. At its simplest, a 3D model can be viewed using a standard desktop computer in real time or as an animation (see 3Deep Media [2012] for a number of examples). The use of experimental or high-end hardware is still relatively rare even though some pioneering projects have emerged during the last decade. Sanders [2011] provides an overview of various maritime archaeology projects that have utilized immersive technology. It should be noted that the term "virtual reality" is often used rather loosely to represent almost anything that contains 3D graphics.

One fundamental factor when comparing different virtual reconstructions of archaeological sites and artifacts is data retrieval. At one end of the spectrum are well-preserved sites that can be accurately measured, photographed, and reconstructed based on their current physical state and particular location. At the other end, researchers may need to base their work purely on representations, such as literature and maps, as well as cope with situations where the physical heritage has been obliterated and what remains is intangible heritage preserved through narratives or rituals. Most real cases fall somewhere between the two extremes. For example, in the case of *Vrouw Maria*, the wreck still exists (albeit at a hard-to-reach location), whereas in an earlier project of ours, the *Finnish Pavilion of 1900*, the building in question was completely demolished over 100 years ago, and the work was mainly based on black-and-white photos. For an in-depth description of the *Pavilion* project, see Díaz et al. [2010].

3. PEOPLE AND PROCESSES

The team that was responsible for building the installation was highly cross-disciplinary. The roles of the group members were not strictly defined, but because of their backgrounds and responsibilities, one half could be labeled as designers and the other as domain experts. The designer half consisted of interaction, visual, sound, and software designers, whereas the domain experts represented maritime archaeology and museology. All in all, the hierarchy was intentionally kept flat in order to foster a creative atmosphere and to create a shared feeling of ownership: people willing to participate were allowed to do so at any point regardless of their background.

As an important part of the project, we held a seminar dealing with various related topics. These types of exchange activities provide opportunities to open up a design space for collaborative engagement and communication between the different knowledge workers. They also serve the purpose of "filling the gaps"—designers needed to understand the basics of maritime archaeology and the history of the wreck, whereas the domain experts were not familiar with digital cultural heritage and interaction design or the possibilities of virtual reality. The lectures, which were of a multidisciplinary nature, included topics such as digital cultural heritage and interface design, the role of memory in the building of a national narrative, the historical background of the Vrouw Maria, interactive narrative development, underwater soundscapes, data acquisition in maritime archaeology, immersive environments and accessibility, and interior architecture and exhibition design.

3.1 Coordinating the Project: Process Models

Designing is a complex, future-oriented activity oriented toward a not-yet-existent but desired outcome. According to Löwgren and Stoltermann [2004], a design project must be "designed itself and depends on creative and innovative thinking for its success." Heritage institutions typically intersect with a wide variety of constituencies and communities in society, which both affords opportunities and poses constraints.

Discussion about the requirements started already during the project negotiation phase. A list with provisions for evaluation of almost everything except the user interface design was drafted. For example, the client specified what items of the landscape and the soundscape should be re-created, as well as the visual style used for rendering the scenes. The production time for the project was 1 year and included two main stages with client review and evaluation meetings every 2 weeks.

From a software development perspective, the overall design and implementation of the installation largely followed the well-known *waterfall model* defined by Winston W. Royce [1970] in his article "Managing the Development of Large Software Systems." The stages originally defined by Royce are system requirements, software requirements, analysis, program design, coding, testing, and deployment. Later on, modified versions have been introduced, with a smaller number of stages or different labeling, but the main concept of sequentially proceeding development has remained the same. As noted already by Royce, the waterfall model is rigid by nature and mistakes made at the beginning will propagate to the later stages, possibly requiring costly and time-consuming backtracking.

The process and methods were also heavily based on the interaction design paradigm as defined by Preece et al. [2002]. As one possibility, they discuss how interaction design can be combined with the waterfall model of software design: roughly speaking, after the requirements have been established, a progressive prototyping and implementation phase follows, followed by a usability evaluation. Most of the design and evaluation methods we employed, such as scenarios, ideation, and formal user tests, are an established part of the standard interaction design toolbox, but due to the experimental nature of the project, we also needed to employ others, such as bodystorming (see following sections).

In our case, it was not possible to define all the interaction design requirements early on, since finding the optimal mode of interaction required extensive iterative testing. Therefore, the software design and implementation phase followed a different model, namely, the *spiral model* introduced by Boehm [1986]. The spiral model is based on cycles where a prototype is constantly improved and redesigned based on the insights provided by reviews and testing. The roots of such iterative—these days often called "agile"—development go back as far as the 1950s [Larman and Basili 2003]. We shall return to the iterative prototyping of the user interface later.

4. CONCEPT DEVELOPMENT

At the beginning of the design process, it was necessary to quickly build shared understanding and terminology for the team. We approached the challenge by progressively building a large *concept map* that helped us chart and define the different facets of the installation-to-be, such as the history of the wreck, underwater conditions, stakeholders involved in the preservation efforts, and the possibilities offered by immersive installations.

4.1 Ideation Workshops: Concept Map, Scenarios, and Information Architecture

Concept maps are based on constructivist learning theories and were originally developed in the 1970s for didactic purposes [Novak and Cañas 2006]. Initially, each member of the group started with his or her individual concepts, which were written on Post-it notes and collected on a whiteboard, where similar notes were clustered to form the main categories of the concept map. Starting individually is useful, since it ensures that all the voices get heard and that the workgroup does not arrive at a consensus too early. The main themes that emerged from the work with the concept map were:

—Experience (how it is to dive)
—Interactive installation
—Nature
—Navigation
—Politics
—Preservation efforts
—Soundscape
—Time
—Treasure
—Users
—Wreck site

We also employed *scenarios* for envisioning the user, his or her needs, and the eventual system design. These scenarios were cocreated with the archaeologists as part of a 2-day ideation workshop. John M. Carroll [1999], a pioneer and proponent of scenario-based design, describes scenarios as stories about people and their activities. The fundamental ingredients of scenarios are the setting (or context), actors, objectives, actions, and events. According to Carroll, the method facilitates reflection about design issues and supports collaboration between the different stakeholders involved in the design process. In addition, scenarios are easy to revise and may cover a wide variety of perspectives to the design problem at hand [Carroll 1999].

Fig. 1. Acting out the scenarios: “three foreign tourists” visiting a museum.

Another property that makes scenarios a lucrative approach is that they can be quick to create and yield useful results before sizeable investments and commitments in a particular technology have been made. Jakob Nielsen discusses scenarios in the scope of discount usability engineering, stating that they help a usability engineer (or in this case, a designer) to focus his or her efforts on the most relevant functionality instead of trying to cover the whole spectrum of possible interactions with a system. In addition to the aforementioned strengths, he suggests that scenarios may even support heuristic evaluation, prototyping, and user testing of a system later on [Nielsen 1995]. Such an observation extends the range of scenario-based design, which is commonly employed only in the first stages of a design process.

During the initial discussions, we recognized three different kinds of user groups that should be targeted in order to cover the spectrum of typical museumgoers: foreign visitors, marine enthusiasts, and schoolchildren who visit the museum with their teacher. Two scenarios were created for each group. Initially, the scenarios were written as one-page text documents that described the interactions with the imagined installation. These descriptions were developed into role-playing performances where three persons improvised the dialogue and acted out the events to the others, who observed and documented the play for further analysis (see Figure 1). Such performances, sometimes known as *informance*, have been used to support collaboration and dialogue between different stakeholders, such as designers and end-users [Burns et al. 1994; Howard et al. 2002; Iacucci et al. 2002]. In our case, indeed, role-playing and informance provided valuable insight to the needs of the three user groups: how to lure people in, how groups differ from individual users, how the physical space needs to be designed, and what types of information should be available.

In addition to enabling us to build shared understanding through the use of concept maps and providing a clearer vision about the users of the installation via the use of scenarios and *personas* [Cooper 1998], the ideation workshop allowed us to create a unique dialogue space where the archaeologists and the designers exchanged information. From the knowledge gained, we designed an information architecture in the form of a 3D water column where we indicated the placement of the wreck in the underwater landscape. *Information architecture* is a term used to describe "the combination of organization, labeling, and navigation schemes within an information system" and the structural design of an information space in order "to facilitate task completion and intuitive access to content" [Morville and Rosenfeld 1998]. This organizational structure was utilized throughout the development as a means of mapping and spatially organizing all the elements in the installation.

4.2 Last Stages of Concept Development

A narrative (or production script) was created in collaboration with the archaeologists. The narrative was based on the knowledge produced by the archaeologists since the research of the wreck began in 2001. It described the contents of the environment and included important historical milestones, such as the fateful storm and the sinking and the rediscovery of the ship, and was like a production script in that it divided into scenes what the users would experience when entering the installation. The observations gained in the scenario workshops and the narrative were used to form a list of defined requirements that the installation should meet.

Using the narrative as a starting point, a visual artist created *storyboards*, which further supported the ideation and bridged the gap between design and implementation. Storyboards have traditionally been used in film production for the early sketching of camera views, visual composition, and action, but more recently they have also found their way to user-centered design, where they are often employed in conjunction with scenarios [Preece et al. 2002]. In our case, their greatest contribution was as boundary objects, enabling the designers and archaeologists to discuss different camera views and their content in concrete terms (cf. [Díaz et al. 2009]).

As the final stage of the concept development, important locations and historical events were compiled into an *interaction matrix*, which described in depth what the user would experience throughout his or her engagement with the installation. The interaction matrix consisted of information broken down into scenes and episodes that consisted of the following properties:

—General description of the location
—Possibilities for user interaction
—User view
—Available extra content, such as videos and photos
—Duration (in the case of chronological events)
—Transition to the following event or location

Figure 2 provides an example of a final info spot showing photographs of the seabed. Similar spots have also been used in other shipwreck reconstructions (e.g., [3Deep Media 2012]). Building the matrix was time-consuming, but it helped to plan the interaction systematically and let us move from concept development to the next phase involving software design and content creation.

5. CONTENT CREATION AND PROTOTYPING

The creation of an immersive installation requires extensive content creation, such as 3D modeling, video/image editing, visual design, sound design, and software development. In the case of the 3D model, we had to combine multiple data sources to build a virtual replica of the site: the seabed was generated from actual depth measurements obtained by sonar scanning, the shape of the hull was based on a laser-scanned miniature model that had been built earlier, and the missing pieces were modeled manually in *3ds Max* using photos, drawings, and video clips as reference material (cf. [Sanders 2011]). Info spots documenting different points of interest were augmented with underwater

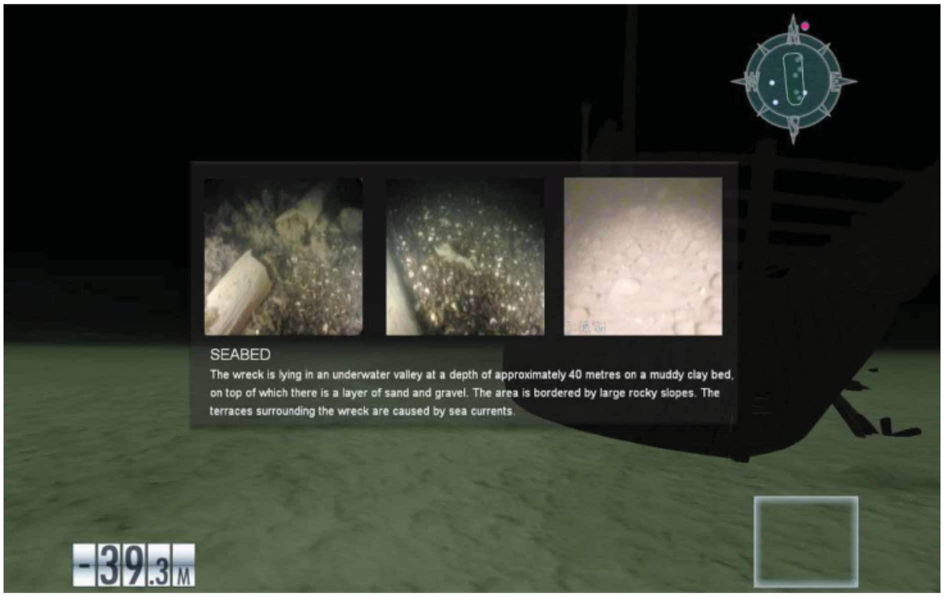

Fig. 2. An info spot describing the seabed near the wreck. Seabed photos courtesy of Peik Joutsen.

footage shot by scuba divers. The design and implementation of the underwater soundscape has been discussed in a separate article [Maras 2012].

5.1 Technical Implementation

Different digital contents were combined in *Unity*, a popular game authoring tool that allows for rapid creation of interactive 3D applications. We went somewhat beyond what is usually done with a game engine, but in the end, *Unity* turned out to be flexible enough for our needs, though special adaptations and external plugins were required for several features. Bringing in the models, textures, and sounds created by the artists was fairly straightforward and little experimentation was needed for that. However, considerable effort was needed to create the gesture-based interaction, stereoscopic 3D using two projectors, and various visual effects such as the ocean, weather, and the fish.

One area that required a significant amount of work was keeping the visual fidelity of the content high, both above and in the sea. One contributing factor is that recent "photorealistic" computer games and CGI-enhanced 3D movies have increased the expectations of the audience. We used two external *Unity* plugins, the *Ocean* and *Unisky* plugins, to create the movement of the water and the weather/sky effects, respectively. They required considerable adjustment on the code level before the result looked and felt as we needed it to. In addition, the fish that swim around the wreck need to be mentioned. We used a simple flocking behavior AI, based on the *Boids* algorithm [Reynolds 1987], to steer them and make their movement look natural. The final version of the 3D graphics can be seen in Figure 3.

Choosing a display device for a VR installation is a question of various tradeoffs between factors such as availability, cost, durability, fidelity, and mobility. We chose a passive back-projected stereo screen mainly because of reliable and affordable passive goggles, high degree of immersion, and—to a certain extent—the novelty value that it offers to the visitors. One of the most important factors was the possibility to display spaces and objects in real scale, as if visiting the actual archaeological site. The downsides, compared to a standard monoscopic screen, were the added cost and complexity of the setup, the need for goggles, and the increased risk of simulator sickness (cf. [Kolasinski 1995]).

Fig. 3. A view from the deck. Visible here are the info spot icons, miniature map, depth meter, and user silhouette viewer (used for testing).

The 3D effect is created via typical linear polarization, using polarizing filters in front of the projector lenses and 3D goggles. In an earlier setup, we had utilized circular polarization, which is not disturbed by the user tilting his or her head, unlike linear polarization. On the other hand, circular polarization suffers from poorer separation between the two images, which causes so-called bleeding in high-contrast areas. The third common technique, known as *active stereo*, is based on shutter glasses that sequentially show half of the frames for each eye. Active stereo does not require a special screen, but the fragility of the relatively expensive goggles and the need to synchronize them with the screen refresh with an infrared beam are major drawbacks. *Autostereoscopic displays*, such as the one popularized by the *Nintendo 3DS* handheld console, do not require tedious goggles and can support multiple users, but their current cost and relatively small size made them unsuitable for our needs. For an overview of immersive display technologies, see Bowman et al. [2005].

5.2 Gestural User Interface

During the ideation sessions, we had already come up with initial ideas for navigation around the virtual space. Physical input devices were considered, but while they do provide useful tactile feedback, they also tend to be fragile (especially a concern when designing for public spaces) and unhygienic and require problematic wires. After evaluating the options that would provide for gestural interaction, we fairly quickly settled on *Microsoft Kinect*, due to its low price and good enough camera quality. A considerable amount of research has focused lately on developing Kinect-based projects in various fields of study, including digital cultural heritage [Richards-Rissetto et al. 2012; Pietroni and Rufa 2012]. During the course of the project, we considered several third-party libraries, eventually settling on *OpenNI* and *NiTE*, that together handle reading the video stream from the device and build a skeletal model out of it. Owing to the limitations of the software and the added complexity of multiple users simultaneously controlling the navigation, the interaction had to be limited to one person at a time. A similar approach has been used, for example, in the *Etruscanning 3D* project, where the Regolini Galassi tomb was reconstructed in VR [Hupperetz et al. 2012].

The first tests of the gestural interface were conducted in a *Wizard of Oz* fashion with an early simple 3D model. Wizard of Oz (WOz) is a term coined by Kelley [1983] and refers to the simulation of a system by a human "computer." For example, systems that rely on natural language for input can be prototyped early on without employing speech recognition software. In this case, an observer interpreted the gestures of the test user and navigated around the model accordingly using a mouse and a keyboard. The test was quick to conduct and produced useful results, as we soon realized that the planned complex sets of gestures were not feasible, since they required learning and were unintuitive. Another metaphor that we evaluated was swimming, which was regarded as engaging but at the same time socially awkward in front of an audience. Furthermore, the test users demonstrated such different ways of swimming that they would have been tedious to support in a coherent manner. Based on those findings, it soon became evident that the most suitable metaphor for navigation was pointing.

Gestural interfaces, including pointing, were pioneered by Myron W. Krueger whose *Videoplace* experiments date back to the mid-1970s [Krueger 1991]. Another milestone of the field is Richard A. Bolt's [1980] influential paper *Put-That-There*. However, "pointing" is not well defined even in the narrow scope of navigation: should the user point at objects or steer by pointing, and what gestures exactly are considered pointing by the system? As noted by Norman [2010], it is important to realize that a gestural interface is not somehow automatically "natural" to users. Half a year of incremental prototyping and testing led us to the following design decisions considering the navigation:

—*Steering by pointing*. Straightforward pointing at objects of interest seemed like an interesting choice at first, but it did not encourage free-form exploration. Free movement is especially important around the wreck where there is plenty of open space.

—*Nonisomorphic rotation*. A nonisomorphic mapping [Poupyrev et al. 2000] is used for turning left and right: the rotation speed is incremented progressively when pointing further away from the midpoint of the screen to allow for both quick turns and steady steering forward.

—*Free exploration*. Instead of taking users semiautomatically through waypoints, we decided to give them the freedom to experience the content in any order. Thus, there could be no predefined linear narrative after the initial descent.

—*Landmark support for navigation*. The info spots were marked with large icons that stood out from the dark environment (cf. [Vinson 1999]).

—*Map support for navigation*. A dynamically updated miniature map was placed in the corner of the screen to help the users orient themselves and find their way back to the wreck if lost in the surroundings.

Even minuscule changes to parameters, such as rotation or travel speed, could lead to significant differences in the overall user experience, so we needed to pay attention to getting the details right. The development of the user interface is discussed in more detail in another article [Sen et al. 2012].

6. USABILITY EVALUATION

A final *user test* was conducted at the Finnish National Museum when the deployment of the installation was getting close, in order to assess how "real" users would perceive it and to fix the remaining usability issues. Two important decisions that had to be made based on the evaluation were the navigation speed and whether we would use a dark or light version of the scene. The backbone of the 2-day session was a test plan consisting of a schedule, six test tasks, questions for the interview, descriptions of the equipment, and guidelines on how to analyze the results afterward.

Fig. 4. A test user trying out the installation.

The test setup was built as closely as possible to the final one (see Figure 4). A stereoscopic screen based on linear polarization was used for displaying the visuals. The sound was only played back using a pair of loudspeakers instead of the final 5.1 surround setup, which was acceptable, since the focus of the test was on the interaction and the immersive graphics. Users navigated around the wreck using the Kinect-based gestural interface described in the previous section. For each test, there was a leader whose main responsibility was to give the user six predefined test tasks and oversee the situation. In addition, at least two observers videotaped the situation and took notes and photos for further analysis. After the tasks, a *semistructured interview* was conducted in order to collect further qualitative feedback on the overall impressions of the test users, the navigation, and the content that they had just seen (cf. [Preece et al. 2002]). All in all, there were 15 test rounds with 23 people—in six cases there was a pair or group of three, so that we could also observe how the installation works with a group.

6.1 Evaluating 3D User Interfaces

A user test is a well-known evaluation method with its established conventions involving a test leader (also known as facilitator), observers, test tasks, videotaping, and so forth. We chose to conduct a typical "lab test" in order to cover the available functionality evenly. Less intrusive user observation would have been less useful in this case, since it could not have guaranteed equal breadth. In the case of virtual environments, there are additional factors that need to be considered when conducting such tests. User tests were chiefly developed for the needs of desktop software, which is in many ways different from immersive installations, where the user may move freely around and use unconventional input and output devices. The concise overview provided by Bowman et al. [2005] mentions the following special characteristics of 3D user interface (3D UI) testing:

—Physical environment issues
—Evaluator issues

—User issues
—Evaluation type issues
—Miscellaneous issues

Several of these issues above could be observed in our test situation too. A projection-based screen typically requires a dark setting, which is adverse to photography and user observation. Videotaping and taking photos of a user in front of a stereoscopic screen produced low-quality images where left and right eye images can be seen on top of each other. The dual-image problem could mostly be solved with a polarized filter in front of the camera lens, but it would make the photos even darker. Due to the imperfect nature of the filtering, there is bound to be some bleeding between the two images in high-contrast areas.

The software, especially the experimental Kinect driver, was not robust enough at this point, which caused a constant need to restart it between and even during the tests. The use of a camera as the input method was the source of other inconveniences as well: low camera placement almost made it impossible for one tall user to participate in the test at all. The clothes of another user confused the tracking so heavily that he was not able to navigate around the wreck until changing to another shirt. Simulator sickness did not hurt the testing significantly, most likely because the navigation speed was intentionally kept low at all times (cf. [Kolasinski 1995]). Four users did, however, report mild headache or eye strain after the test.

6.2. Observations and Their Use

The analysis of the observations made during the tests progressed in a typical manner: individual observations were grouped, described in detail, and subjectively ranked according to their severity on a scale from zero to three (note, cosmetic, medium, fatal). The two main factors considered in the severity rating were the frequency of the issue and its effect on the interaction. Usability studies are often focused on finding errors, which is, naturally, important, but positive findings are equally useful. Therefore, we also collected positive examples that highlighted parts of the installation that were easy to use and worked well. The test tasks and their completion rates were as follows (out of 15):

(1) Navigate to the rear of the ship (14)
(2) Find main cargo bay (11)
(3) Discover cargo bay contents and exit (5)
(4) Find information on the masts (12)
(5) Go to the sea bed (12)
(6) Return to the ship (8)

Finally, the findings were turned into suggestions for improvements that were prioritized so that at least the most severe problems could be fixed in the remaining time. The problems rated three (fatal) were:

—Initial calibration and instructions incomprehensible
—Unreliable gesture tracking
—Problems with vertical navigation
—Problems with stopping
—Exiting the cargo hold difficult
—Getting lost when far from the ship

Some of the most imminent problems were related to the initial calibration that was required before starting the use: users would not understand what they were expected to do or, when they could start the navigation, problems that were accentuated by the tracking system, which was not completely reliable at that point. These problems were mostly solved by refining the gesture detection and improving the on-screen instructions. The navigation itself was not error free either: people had difficulty stopping the motion and steering up or down. In addition, it was hard to hit the info spots without running past or through them. Again, tweaking the parameters of the gestural interface was needed. Pointing around for 10 minutes or more with the arm extended started to get tiring—something that cannot be completely solved if the input is so heavily based on pointing.

Certain issues had more to do with the content or the presentation than the gestural interface. In particular, getting back to the deck from the cargo hold was problematic for practically everybody, since you could only exit through the hatches that were hard to find. Likewise, entering and exiting the ship was troublesome for several users: they could have simply steered over the railing but were unwilling to do so. Some users got lost wandering on the seabed around the wreck and could not find their way back, even if the miniature map was there exactly for that purpose. Based on the interviews, it seems that people hardly noticed the map or the depth meter (see Figure 2), because they were so focused on the actual 3D content. The lengthy introduction that shows the sinking and deterioration of the ship had to be cut significantly shorter, since many users wanted to start navigating around immediately instead of passively waiting for their turn.

On the positive side, all except one test user found the installation enjoyable to use and the content interesting. One more observation was that previous—in some cases extensive—knowledge of the wreck let users experience the installation on a significantly deeper level than mere exploration. The worst problems were caused by the unfinished implementation of the gestural interface, which was reworked before the final deployment of the installation. As to the original questions about the speed and darkness, we decided to keep the navigation speed low and enhance the contrast of the scene: on the one hand, darkness was considered more atmospheric than a well-lit view, but on the other, users wanted to see more interesting details that would have disappeared if the whole scene were darker. The visibility on the actual wreck might be as little as 1 meter during the worst season.

7. CONCLUSION

We have presented here one possible approach to the user-centered design and implementation of immersive cultural heritage installations, particularly in the context of gesture-based interaction and maritime archaeology. All in all, the methods we employed throughout the project were found useful: they provided insight, inspiration, and findings that could be utilized at later stages. The spectrum of methods covered here is already considerably wide for a single project, but there are several others that might prove equally useful. For example, *focus groups* (sessions where actual end-users discuss and work with designers) might have provided additional insight to the needs and attitudes of the museumgoers. Different types of *user observation* could have helped the designers better understand people's activities in a museum setting. One more notable method is *use cases* (detailed descriptions of interaction taking place between a user and a system), which serve the same purpose as interaction matrices that were discussed in Section 4.2. (cf. [Preece et al. 2002])

The relatively tight schedule of the project (proceeding along other curricular activities at a university) is important when considering the value of the methods used. In this case, concept mapping, scenarios (combined with role-playing and informance), and Wizard of Oz prototyping produced useful results quickly and economically, owing to their lightweight nature. As an example, a session that consisted of writing, acting out, and documenting six scenarios was carried out in just 1 day by a group of nine participants. Interaction matrices, user tests, and especially iterative prototyping represent the other end of the scale: they also supported the process but were considerably heavier to deploy. Storyboards fall somewhere between the two extremes. A relevant observation on storyboarding was that it requires careful planning in order to serve any real purpose.

There were notable differences between the methods in their applicability to immersive installations. The use of experimental input and output devices hardly affected the concept development phase at all, whereas prototyping and user testing were markedly different to desktop software development. To summarize: based on our experiences, the concepts, needs, and requirements are all about users, not technology, which makes it possible to tackle very different design tasks using the same initial approach. The need to consider the particularities of the target medium arises only at later stages of the design process. In relation to this, the role of domain experts was emphasized early on and became less pronounced toward the implementation. Another way to evaluate the methods is the generalizability of their outcomes. In this case, the concept maps, storyboards, scenarios, and interaction matrices we created were, in practice, specific to this particular project and will only be of use as reference material in the future. In comparison, the lessons learned during the interactive prototyping of the gestural interface are directly applicable to other similar projects. As a matter of fact, the user interface for another installation was already ported over from the codebase of *Vrouw Maria* with little effort.

All in all, the end result, an immersive digital cultural heritage installation, works well after the corrections and fulfills the initial requirements. As is always the case, some tradeoffs had to be made in order to improve the most important aspects of the system. For example, with a camera-based interface, it is hard to achieve the same coherence and accuracy as with a physical controller, but on the other hand, the use of natural gestures, the absence of wires, and durability (especially of concern in a public space) are undeniable strengths of such an approach. After evaluating a wide range of different possible gesture sets, it became evident how navigating by pointing is practically the only metaphor that does not require extra learning (again, of essence in a public space with constantly changing users). Even though "pointing" might appear trivial at first, it is considerably difficult to get it to work fluently, as differences between users and their gestures, as well as various parameters ranging from acceleration to rotation speed, need to be taken into account.

The main focus of the discussion here has been on the design and implementation side of things. However, it is hard to overstate the importance of maritime archaeology throughout the process. The role of historical research goes way beyond mere content, such as textual descriptions, photographs, and video footage. The combination of archaeology and design was the backbone of the project: it provided the team with continuous inspiration and rich narratives, as well as practical means to develop and implement a shared vision.

ACKNOWLEDGMENTS

We thank our collaborators at the Finnish National Board of Antiquities and the Maritime Museum of Kotka for making the *Vrouw Maria* installation possible, and all the Media Lab Helsinki students who contributed to the project.

REFERENCES

3Deep Media. 2012. Digital 3D Modelling of Shipwrecks and Other Underwater Sites. Retrieved May 29, 2014, from http://3deepmedia.co.uk/services/3d-shipwreck-modelling.php.

Elizabeth A. Bartley and John E. Hancock. 2008. Virtual reconstructions as destination tourism? *International Journal of Digital Culture and Electronic Tourism* 1, 2–3 (2008), 225–239.

Barry Boehm. 1986. A spiral model of software development and enhancement. *ACM SIGSOFT Software Engineering Notes* 11, 4 (1986), 14–24.

Richard A. Bolt. 1980. Put-That-There: Voice and gesture at the graphics interface. In *Proceedings of SIGGRAPH'80*. ACM, New York, 262–270.

Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2005. *3D User Interfaces: Theory and Practice*. Pearson Education, Boston, MA.

Colin Burns, Eric Dishman, William Verplank, and Bud Lassiter. 1994. Actors, hair-dos and videotape: Informance design. In *Proceedings of CHI'94*. ACM, New York, 119–120.

John M. Carroll. 1999. Five reasons for scenario-based design. In *Proceedings of the 32nd Hawaii International Conference on System Sciences*. IEEE, Washington DC, 3051–3061.

Alan Cooper. 1998. *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity*. Macmillan Computer Publishing, London.

Lily Díaz, Markku Reunanen, and Anna Salmi. 2009. Role playing and collaborative scenario design development. In *Proceedings of the 17th International Conference on Engineering Design (ICED'09)*. Design Society, Castle Cary, 79–86.

Lily Díaz, Markku Reunanen, and Anna Salmi. 2010. Digital reconstruction of the Finnish 1900 world fair pavilion: Experiences from creating a virtual reality installation. In *Heritage in the Digital Era*, Marinos Ioannides, Addison Alonzo, Andreas Georgopoulos, Loukas Kalisperis, Andre Brown, and Denis Pitzalis (Eds.). Multi-Science Publishing, Essex, 167–180.

Eero Ehanti, Johanna Aartomaa, Irma Lounatvuori, and Erik Tirkkonen (Eds.). 2012. *Lost at Sea, Rediscovered*. National Board of Antiquities, Helsinki.

F.A.B.R.I.CATORS. 2011. The Last Supper Interactive. Interactive Narrative Embedded in Immersive Virtual Environments. Retrieved May 29, 2014, from http://www.fabricat.com/PDF_FILES/Last Supper Interactive_FSAS_Fishcnaller.pdf.

Steve Howard, Jennie Carroll, John Murhpy, and Jane Peck. 2002. Using 'Endowed Props' in scenario-based design. In *Proceedings of the 2nd Nordic Conference on Human–Computer Interaction (NordiCHI'02)*. ACM, New York, 1–10.

Wim Hupperetz, Raffaele Carlani, Daniel Pletinckx, and Eva Pietroni. 2012. Etruscanning 3D project. The 3D reconstruction of the Regolini Galassi tomb as a research tool and a new approach in storytelling. *Virtual Archaeology Review* 3, 7 (2012), 92–96.

Giulio Iacucci, Carlo Iacucci, and Kari Kuutti. 2002. Imagining and experiencing in design, the role of performances. In *Proceedings of the 2nd Nordic Conference on Human–Computer Interaction (NordiCHI'02)*. ACM, New York, 167–176.

John F. Kelley. 1983. An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of CHI'83*. ACM, New York, 193–196.

Eugenia M. Kolasinski. 1995. *Simulator Sickness in Virtual Environments*. Technical Report 1027. US Army Research Institute for Behavioral and Social Sciences, Alexandria, VA.

Myron W. Krueger. 1991. *Artificial Reality II* (2nd ed.). Addison-Wesley, Boston, MA.

Craig Larman and Victor R. Basili. 2003. Iterative and incremental development: A brief history. *IEEE Computer* 36, 6 (2003), 47–56.

Jonas Löwgren and Erik Stoltermann. 2004. *Thoughtful Interaction Design: A Design Perspective on Information Technology*. MIT Press, Cambridge, MA.

Svetlana Maras. 2012. Recreating the underwater soundscape in the Vrouw Maria virtual simulation. In *Lost at Sea, Rediscovered*, Eero Ehanti, Johanna Aartomaa, Irma Lounatvuori, and Erik Tirkkonen (Eds.). National Board of Antiquities, Helsinki, 156–163.

Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1994. Augmented reality: A class of displays on the reality-virtuality continuum. In *Proceedings of SPIE 2351, Telemanipulator and Telepresence Technologies*. Hari Das, Boston, MA, 282–292.

Peter Morville and Louis Rosenfeld. 1998. *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*. O'Reilly Media, Sebastopol, CA.

Jakob Nielsen. 1995. Scenarios in discount usability engineering. In *Scenario-Based Design: Envisioning Work and Technology in System Development*, John M. Carroll (Ed.). John Wiley & Sons, Hoboken, NJ, 59–83.

Donald A. Norman. 2010. Natural user interfaces are not natural. *Interactions* 17, 3 (2010), 6–10.

Joseph D. Novak and Alberto J. Cañas. 2006. *The Theory Underlying Concept Maps and How to Construct and Use Them*. Technical Report. Institute for Human and Machine Cognition, Pensacola. Retrieved June 24, 2013, from http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlying ConceptMaps.htm.

Eva Pietroni and Claudio Rufa. 2012. Natural interaction in virtual environments for cultural heritage: Giotto in 3D and etruscanning study cases. In *Virtual Systems and Multimedia (VSMM) 2012*, Gabriele Guidi and Alonzo C. Addison (Eds.). IEEE, New York, 339–346.

Place-Hampi. 2008. PLACE-HAMPI :: Inhabiting the Cultural Imaginary. Retrieved July 9, 2013, from http://www.place-hampi.museum/contents/place-hampi.html.

Ivan Poupyrev, Suzanne Weghorst, and Sidney Fels. 2000. Non-isomorphic 3D rotational techniques. In *Proceedings of CHI'00*. ACM, New York, 540–547.

Jennifer Preece, Yvonne Rogers, and Helen Sharp. 2002. *Interaction Design: Beyond Human–Computer Interaction*. John Wiley & Sons, Hoboken, NJ.

Craig Reynolds. 1987. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of SIGGRAPH'87*. ACM, New York, 25–34.

Heather Richards-Rissetto, Fabio Remondino, Giorgio Agugiaro, Jennifer von Schwerin, Jim Robertsson, and Gabrio Girardi. 2012. Kinect and 3D GIS in archaeology. In *Virtual Systems and Multimedia (VSMM) 2012*, Gabriele Guidi and Alonzo C. Addison (Eds.). IEEE, New York, 331–337.

Winston Royce. 1970. Managing the development of large software systems. In *Proceedings of IEEE WESCON*, Vol. 26. IEEE Computer Society, Washington DC, 1–9.

Donald H. Sanders. 2011. Virtual reconstruction of maritime sites and artifacts. In *The Oxford Handbook of Maritime Archaeology*, Alexis Catsambis, Ben Ford, and Donny L. Hamilton (Eds.). Oxford University Press, Oxford, 305–326.

Ferhat Sen, Lily Díaz, and Tommi Horttana. 2012. A novel gesture-based interface for a VR simulation: Re-discovering Vrouw Maria. In *Proceedings of Virtual Systems and Multimedia (VSMM) 2012*. IEEE, New York, 323–330.

V-Must. 2011. Virtual Museum Transnational Network, v-must. Retrieved November 25, 2013, from http://www.v-must.net/virtual-museums.

Norman G. Vinson. 1999. Design guidelines for landmarks to support navigation in virtual environments. In *Proceedings of CHI'99*. ACM, New York, 278–285.

Vassilios Vlahakis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Luis Almeida, Didier Stricker, Tim Gleue, Ioannis T. Christou, Renzo Carlucci, and Nikos Ioannidis. 2001. Archeoguide: First results of an augmented reality, mobile computing system in cultural heritage sites. In *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage (VAST'01)*. ACM, New York, 131–140.

Kathleen Wilson. 1988. Palenque: An interactive multimedia digital video interactive prototype for children. In *Proceedings of CHI'88*. ACM, New York, 275–279.

Received July 2013; revised May 2014; accepted June 2014