

Virtual Valcamonica: Collaborative Exploration of Prehistoric Petroglyphs and Their Surrounding Environment in Multi-User Virtual Reality

Alexander Kulik, André Kunert, Stephan Beck, Carl-Feofan Matthes, Andre Schollmeyer, Adrian Kreskowski, Sue Cobb, Mirabelle D’Cruz, Bernd Froehlich

Virtual Reality and Visualization Research Group at Bauhaus-Universität Weimar, Human Factors Research Group, University of Nottingham, UK

ABSTRACT

In this paper, we present a novel, multi-user, virtual reality environment for the interactive, collaborative 3D analysis of large 3D scans and the technical advancements that were necessary to build it: a multi-view rendering system for large 3D point clouds, a suitable display infrastructure and a suite of collaborative 3D interaction techniques. The cultural heritage site of Valcamonica in Italy with its large collection of prehistoric rock-art served as an exemplary use case for evaluation. The results show that our output-sensitive level-of-detail rendering system is capable of visualizing a 3D dataset with an aggregate size of more than 14 billion points at interactive frame rates.

The system design in this exemplar application results from close exchange with a small group of potential users: archaeologists with expertise in rock-art and allows them to explore the prehistoric art and its spatial context with highly realistic appearance. A set of dedicated interaction techniques was developed to facilitate collaborative visual analysis. A multi-display workspace supports the immediate comparison of geographically distributed artifacts. An expert review of the final demonstrator confirmed the potential for added value in rock-art research and the usability of our collaborative interaction techniques.

1 INTRODUCTION

The advancements of 3D scanning technologies over the last decade have made it possible to acquire highly precise 3D models of cultural heritage sites and their exhibits. In many cases, however, the visualization of isolated virtual objects captured from a cultural heritage site, lacks important information provided by the surrounding context in which the artifact was created or placed and thus denies the user rich information that would enhance their understanding and interpretation of cultural heritage. We developed the necessary technology and infrastructure to present 3D scanned artifacts and their surrounding as a coherent and freely explorable virtual environment for groups of researchers or museum visitors.

A multi-disciplinary team of researchers comprising archaeologists, computer scientists, engineers, scanning and human factors experts worked together as part of the 3D-Pitoti project to address the challenges of acquisition, processing and presentation of 3D archaeological data, focusing on pre-historic rock art in Valcamonica, Italy¹. This paper describes a novel, multi-user, virtual reality (VR) system that allows users to collaboratively explore and analyze large 3D scans of prehistoric rock-art and their environmental context (Figure 1). This includes a novel real-time rendering technique for the visualization of large 3D point-cloud models in multi-user virtual reality. Our point-based rendering system supports the interactive visualization of cultural heritage sites and their surroundings with hundreds of gigabytes or even terabytes of data,



Figure 1: The 3D-Pitoti Scientists’ Lab involves a large vertical multi-user 3D display (powerwall) and a multi-touch enabled multi-user 3D tabletop display. We use both displays for different subtasks of the data exploration. Handheld virtual 3D displays, called Photoportals, provide additional access points for individual or collaborative interaction with the data.

despite limited resources in terms of bandwidth, video memory and main memory.

Our contributions can be summarized as follows:

- An output-sensitive, real-time rendering system for the interactive presentation of very large 3D point cloud data in collocated multi-user virtual reality.
- A unique multi-user 3D visualization infrastructure consisting of multiple shared displays
- A suite of 3D interaction techniques for the collaborative analysis of large 3D scanning data

Based on these technical developments, we co-designed with archaeologists a collaborative, virtual-reality workspace: the 3D-Pitoti Scientists’ Lab. Our experimental implementation of the envisioned system allows users to immediately share perspectives, test each other’s hypotheses, and present evidence directly in the jointly perceived 3D model. An evaluation was conducted via a qualitative user study with five domain experts and found high levels of reported usability and usefulness of our technology developments for archaeological research.

¹<http://www.3d-pitoti.eu>, http://cordis.europa.eu/result/rcn/182986_en.html

2 RELATED WORK

3D modeling, laser-scanning, LiDAR, and photogrammetry introduced valuable novel survey methods for archaeological research (Chapman et al., 2013; Thwaites, 2013). The digital data enables broad accessibility and offers novel opportunities for analysis (Olson et al., 2013; Soler et al., 2017), preservation (Kim et al., 2006), virtual reconstruction (White, 2013), and simulation (Ch'ng & Gaffney, 2013). Prior work has also shown that interactive displays of such digital 3D models can provide more appealing ways for museum visitors to engage with cultural heritage (Bruno et al., 2010). In particular, Di Franco et al. reported that people appreciate the exploration of ancient artifacts through direct manipulation (Di Franco et al., 2015). Moreover, such virtual museums are interactive archives that offer new ways to accumulate and disseminate scientific knowledge (Kim et al., 2006).

Prominent examples include visualization of landscapes (Chase et al., 2011), lost cities (Henderson et al., 2013), ancient buildings (Pierdicca et al., 2016), prehistoric objects (Grosman et al., 2008), and virtual models of inaccessible artifacts such as sunken ships (Barreau et al., 2015; Reunanen et al., 2015). Domingo et al. showed the benefits of 3D recording techniques in combination with 2D tracings for the documentation of Levantine rock art in Spain (Domingo et al., 2013). The scanning, reconstruction and display of single objects can often be achieved with readily available technologies (Pavlidis et al., 2007; Olson et al., 2013; Webel et al., 2013), but context information is often missing. Bruno et al. therefore suggested to incorporate contextual information through additional videos, pictures, music and 3D animations (Bruno et al., 2010). As an alternative to such a collection of different media we suggest the presentation of detail scans embedded within their environmental context and support for simultaneous presentation of multiple detail scans to facilitate their immediate comparison. This required the development of multi-scale 3D acquisition methods (Alexander et al., 2015) and novel interactive visualization techniques.

3 ROCK ART RESEARCH IN VALCAMONICA

Valcamonica rock-art is spread over an entire valley of about 70 kilometers length in the Lombardy region of Italy (Figure 2). It is the largest collection of prehistoric rock engravings (petroglyphs) in Europe with an estimate of over 300,000 individual artworks on more than 2,000 natural rock surfaces (Smith, 2014, pp. 7583-7587). The time of their creation spans from around 3500BC to the late Middle Ages (Marretta, 2013). More than two-thirds of the petroglyphs are attributed to the Iron Age (Anati, 1961). This wealth of European cultural heritage provides autobiographical stories from Europe's ancient past and its cultural importance was recognized by the award in 1979 of Valcamonica as a UNESCO World Heritage site.

Most of the engravings show abstractions of people (e.g. warriors, farmers, dancers), animals (e.g. deers, dogs, birds), and objects (e.g. maps, letters, crosses, keys) (Figure 3). The locals call them *Pitoti* which means "little puppet" in the Camuni dialect (Chippindale & Baker, 2012). Traditional archaeological methods for capturing this data include drawing, rubbing (frottage), casts and tracing (Anati, 1977). However, these methods of data capture can be time consuming and difficult, also because some rock art panels are located in quite inaccessible locations.

Much of the interpretation of Valcamonica rock art is based on theories of meaning of the pictorial 'stories' represented by the collection of images on individual rock panels. These theories include the development of humanity's religious ideas, sensibilities, and rituals (Anati, 1961, 1976, 1977; Anati & Cittadini, 1994), e.g., the initiation of male aristocratic warriors during the Iron Age and separate male and female ritual spheres during the Bronze Age (A. Fossati, 1991; A. E. Fossati, 2008).

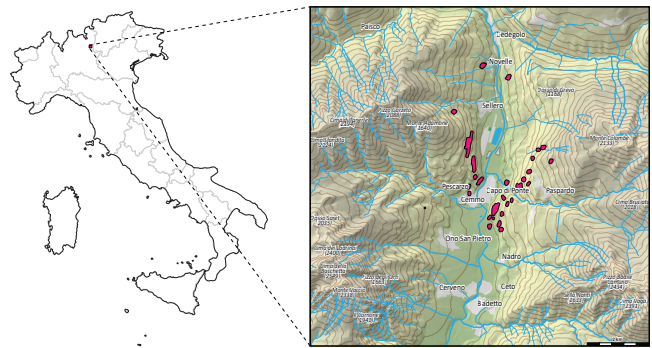


Figure 2: Valcamonica in the Lombardy region of Italy contains the largest collection of prehistoric rock art in Europe. The pink regions in the detail callout indicate major rock art sites (mapping from Alexander, 2011). The rock art panels are surrounded by impressive mountains. Pizzo Badile Camuno (2,434 m) is located in the southeast. La Concarena in the southwest incorporates three summits: Cima della Bacchetta (2,549 m), Cima dei Ladrinai (2,403 m), and Monte Vaccio (2,338 m).

Other interpretations consider the significance of the location, given that the valley is between two mountains of around 2,500 m in height (Pizzo Badile Camuno and La Concarena, Figure 2). Under certain atmospheric conditions, the valley has a "corona effect" which supports the view that it was regarded as a sacred place (Priuli, 1985). More recent research has emphasized the variation in imagery found in different places across the valley. Marretta noted that different types of bird image were found in distinct spatial distributions while Alexander observed that motifs appear in distinct groupings in different locations (Marretta, 2007; Alexander, 2011). Such exploration is important because prior research assumed that all variation in the Pitoti was diachronic (occurring across time) whereas the presence of spatial patterns as described above suggests that at least some variation is also synchronic (occurring at the same time but across space).

However, traditional methods of data capture and cataloguing can make it difficult to identify relevant characteristics and similarities in rock-art imagery occurring in different locations. Tracings show the figures as if they were flat and clearly visible. A photo shows the appearance only for a single viewpoint and the lighting conditions present at the time that the photo was taken. The actual visibility of the petroglyphs, however, is dynamic. Many of the figures can only be seen with incident light from a very shallow angle (e.g. during dusk or dawn) that reveals the small images through cast shadows (Figure 3(c)). This dynamic appearance is potentially a relevant feature which might indicate how and when the Camuni people used the artworks to tell their stories.

Interpretation of the rock engravings based on their visual representation alone fails to take account of contextual features of their environment. For example, it can be expected that the curvature of the rock surfaces played a role in their selection for a particular artwork. Moreover, individual Pitoti differ substantially in their 3D structure as they have been hand carved by different artists, on different rocks, in different periods of time. Furthermore, the petroglyphs are not evenly distributed over all rock panels in the valley. Instead, they are concentrated on a few specific sites. These clusters and their spatial relations can be demonstrated with overview maps and detail views, but it is a very different experience to perceive the ancient art galleries from an egocentric perspective, embedded in the surrounding landscape with the figures pecked into uneven rock panels.



(a)



(b)



(c)

Figure 3: Rock engravings in the Valcamonica valley (a). In many cases, the petroglyphs are barely visible (b) and incident sunlight or artificial light sources (c) are necessary to reveal them.

4 LARGE MULTI-RESOLUTION 3D SCANNING DATA

Our project partners acquired a 3D model of the Valcamonica rock art and its natural environment using various 3D scanning technologies. The model was composed of many individual 3D scans that differ in their size, resolution, surface properties and scanning methods (Alexander et al., 2015). More specifically, the data was acquired at three different scale levels (Figure 4):

Large-scale scans were acquired for the valley overview and cover up to some square kilometers with a resolution of about 10 cm. Structure-from-Motion (SfM) reconstruction methods were applied to derive the model from aerial photography with an ultralight aircraft. The resulting 3D model of the valley serves as a geographical reference for orientation and navigation between the sites of interest.

Medium-scale scans of rock panels typically cover several square meters with a resolution between 2 and 8 mm. They were acquired with terrestrial laser scanners or derived from aerial photography with unmanned aerial vehicles (UAV). The scans of the rock panels enable a realistic appearance of the direct environment of the Pitoti when exploring the ancient artworks at original size.

Small-scale scans of the individual Pitoti figures with an accuracy of about 0.1 mm were captured with structured-light scanning and SfM-methods. The high resolution allows a detailed inspection of the rock art by means of magnification. Such close-up views can reveal geometric features that would be hardly visible for the human eye on site.

The visualization of the Pitoti and their surroundings required the integration of all scanned models into a coherent virtual environment. However, the differently scaled models cannot be simply merged. Ambiguities, where multiple scans overlap, have to be resolved. Moreover, different acquisition techniques and changing lighting conditions between different scans result in color differen-

ces that should be adjusted. In this paper, however, we focus on the visualization and interaction challenges.

5 REQUIREMENTS ANALYSIS

The 3D model of Valcamonica and its rock art enables the visualization and analysis of the prehistoric artifacts at any time and any location. However, the changing context of its presentation affects the viewer's perception and interpretation. We therefore aimed to retain the realistic appearance of the artworks embedded in their natural surroundings. The interactive display system still creates a perceptual context of its own, but at least the actual geographic context can be perceived correctly and taken into consideration for the archaeological analysis and interpretation. Besides this fundamental feature, the technology must also offer relevant interactive functionalities for archaeological research tasks. We believe that this includes support for immediate exchange and collaboration with others. Consequently, our developments were informed by general groupware usability requirements as described in the literature and more specific functional requirements that were gathered from early test users.

5.1 Groupware Requirements

One of the primary requirements for collaborative interaction is mutual awareness among participants. Gutwin and Greenberg defined the concept of workspace awareness with three main aspects: 1. the presence, identity and authorship of participants (who), 2. the involved artifacts, actions and intentions (what), and 3. the location, gaze, view and reach of users (where) (Gutwin & Greenberg, 2002). Their concept also entails the history of artifacts and events (how and when). Besides intentional communication through words and gestures, they also identified more implicit modes of information exchange in real-world settings, e.g., through the observation of each other's activity (consequential communication) and the multi-sensory feedback that results from the manipulation of artefacts (e.g. sounds).

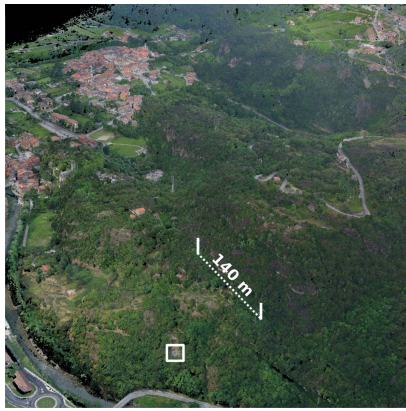
Baker et al. (2002) specified more concrete groupware usability heuristics. Their demand to support intentional verbal and gestural communication, but also consequential communication of the participants' embodiments and their shared artefacts, reflects the concept of workspace awareness. They also emphasized the necessity of enabling continual transitions between tightly and loosely-coupled collaboration as well as protection from mutual interference. In real-world settings, this behavior often results in the emergence of separate territories for private interaction and mutual exchange (Scott et al., 2004). Not least, Baker et al. noted that groupware should support mutual coordination of actions as well as informal encounters and spontaneous collaboration. Besides workspace awareness, collaborative systems should thus offer opportunities to assume complementary roles (Benford et al., 2000).

From this and other related work, we derived three core topics for the design of collaborative systems to comply with the above mentioned requirements and to enable effective cooperative interaction: Workspace Coherence, Emergent Territoriality, and Complementary Capabilities (Kulik, 2016). The spatiotemporal coherence of the shared workspace supports mutual awareness. Support for emergent territoriality facilitates mutual coordination with minimal interference and encourages frequent transitions between tightly and loosely coupled collaboration. A diversity of available tools and interaction techniques with complementary functionality facilitates the coordination of concurrent activities and mutual support.

5.2 Specific User Requirements

To inform our understanding of end-user requirements, w

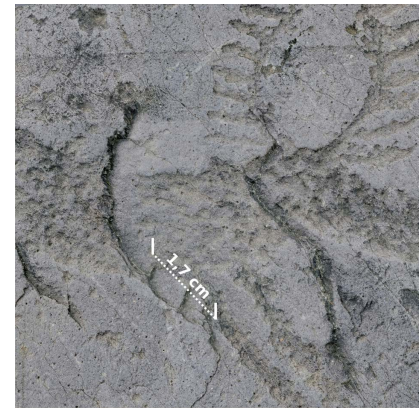
To inform our understanding of end-user requirements, we pursued a user centered design process that involved rock-art archa-



(a) The 3D model of Valcamonica covers a real-world extent of 1.5 kilometers and consists of more than 14 billion scanned points.



(b) The rock-art site Seradina 12C in the valley and detailed scans of some exemplary artworks amount to 4 billion points.



(c) Detail view of a deer figure at Seradina 12C. The petroglyphs are difficult to perceive and analyze, since the peck marks are only a few millimeters in size.

Figure 4: The multi-resolution model of virtual Valcamonica ranges from an overview of the entire valley (a), over medium-scale rocks (b) to detailed scans at submillimeter scale (c). To convey an impression of this multi-scale dataset, the approximate position of (b) and (c) are marked with white squares.

eologists throughout the research project. In the early stages a two-day workshop identified the archaeologists' requirements and how they would want to use the envisioned 3D-Pitoti Scientists' Lab. The participants comprised a team of 12 experts in either rock art and/or landscape archeology: five who were involved in the 3D-Pitoti project and seven external participants. The presented technologies included experimental 3D projection displays for single and multiple users, a 2D multi-touch display and fundamental 3D interaction techniques. The participants provided feedback on each demonstrator and were then asked to discuss how they would want to use such data and technology for their archaeological research. The feedback provided rich information from which two use case scenarios were derived together with a list requirements for the technology developers. Key requirements are discussed below.

5.2.1 Requirements for rock art details

The archaeologists stressed the importance of shared perception of the scanned petroglyphs in order to discuss their findings and viewpoints. They appreciated the possibility of scaling the artworks for a better perception of the geometric details and requested support to frequently adjust both the scaling of, and the viewing distance to, the analyzed data. This included the desire for a virtual magnifying glass, as used in fieldwork to focus on particular features, while keeping the surrounding context. Direct control over the lighting conditions was also mentioned as a requirement together with an explanation of the corresponding fieldwork methods including mirrors, shades and flashlights. The archaeologists also requested the means to directly compare scanned 3D surfaces from different locations in the valley and potentially also from other rock art sites around the world. To similar ends, they wanted the system to include data of other modalities, e.g. photos, tracings and descriptions. The importance of photorealistic appearance was stated, but the archaeologists also asked for more abstract visualization options (including color-coded depth and curvature maps as well as cross section views). Among those, they were particularly fascinated by the possibility to view the surfaces from "inside the rock".

5.2.2 Requirements for large scale environments

The archaeologists stated that the possibility for rapid scaling and view navigation was even more important for larger 3D datasets, e.g. landscapes. However, during the experimental sessions with

the VR demonstrators, several among them commented that they lost their sense of scale during the virtual navigation. A virtual sliding scale was therefore strongly requested to convey the relative size of the displayed data. Two different types of navigation were desired: Continuous motion according a 'flying' metaphor was deemed necessary to experience spatial relations in the virtual environment; and discrete transitions, e.g., to switch between life-size egocentric viewpoints, for example to conduct a viewshed analysis. With regard to the controlled lighting conditions, they requested a sunlight simulation that reflects the actual geographic location and can be specified in terms of date and time.

5.2.3 General Requests

Several listed requirements were common to both use cases or required similar technical developments. Among those, a predominant request of our participants was to have data interfaces to and from the virtual environments for analysis. For example, they wanted to add data from heterogeneous sources for scientific comparison in the immersive setting. In the other direction, they asked for possibilities to export views and insights from the virtual exploration to other media that can be more easily shared with colleagues, e.g. images for inclusion in published articles. Many further requests were directly related to information management and databank access embedded in the virtual environment. Most fundamentally, they wanted to be able to query and display the locations and densities of different artworks. Moreover, access to information related to individual artworks and environmental features was requested. This included information about the types of rock and vegetation but, more importantly, the available scientific information about particular types of figures and their artistic styles. Ideally, machine learning techniques could be used to automatically segment human artwork from natural traces and classify the different styles of their creation. Some participants that were keen on quantitative and statistical methods, asked for the functionalities provided by their established GIS packages.

6 THE 3D-PITOTI SCIENTISTS' LAB

The specified user requirements were incorporated into the 3D-Pitoti Scientists' Lab to provide a prototypical visualization infrastructure for the collaborative exploration and analysis of large and

complex 3D scanning data such as the Valcamonica rock-art. (Figure 1). Its core elements are two multi-user 3D displays, a real-time rendering system for large 3D point clouds, and a suite of 3D interaction techniques for collaborative 3D data analysis.

6.1 Multi-User 3D Displays

We consider the coherence of the collaborative interaction space to be the most fundamental requirement. If all involved users can see themselves as well as everybody else and perceive also the shared virtual environment corresponding to their physical viewpoint, workspace awareness can be achieved based on implicit cues Gutwin & Greenberg (2002). Therefore, we built our setup on multi-user projection technology Kulik et al. (2011). Our projectors are running at 360 Hz such that six different views can be displayed within a 60 Hz frame. Up to six stereo views can be achieved by combining two of these projectors and adding passive channel separation through polarization filters. Our prototype system supports three users with individual stereo views at two synchronized 3D displays, a tabletop display of size 1.13m x 0.85m and a large vertical display (powerwall) of size 4.3m x 2.7m.

6.2 Multiple Displays and Viewing Windows

Our set-up involves two physical screens and additional virtual viewing windows (Figure 1). The devices complement each other's functionality with different affordances and provide many opportunities for users to work partially in private interaction territories inside a shared interaction space. The tabletop display, for example, supports allocentric overviews of a miniaturized virtual environment, while the larger powerwall enables an immersive egocentric perspective of the scene at original scale.

Our virtual viewing windows are 3D portals to different locations in the scene. They can be used alone or by subgroups of users. Our implementation follows the earlier suggested metaphor of photography to capture and show distant locations Kunert et al. (2014). The mobile and dynamic 3D views shown in a Photportal allow users to search individually for interesting perspectives. New navigation targets can then be shown to others and even entered through the portal by the whole group.

7 REAL-TIME RENDERING OF LARGE 3D SCANNING DATA IN MULTI-USER VIRTUAL REALITY

The interactive, stereoscopic visualization of the Valcamonica valley and its rock-art for a group of archaeologists posed several challenges to the underlying rendering system. The 3D scans need to be visualized in real-time for multiple observer perspectives and with an authentic appearance on high-resolution 3D displays. This requires an output-sensitive rendering system which means that only visible parts of the scene should be rendered and the model fidelity should adapt to the current camera view and screen resolution.

Scanned data sets frequently come in the form of point clouds where each sample features a color or a radiometric material component. Commonly, mesh reconstruction approaches (e.g. Kazhdan & Hoppe, 2013; Bradley et al., 2008) are used to represent scanned data sets in a more compact, closed and convenient form. However, these approaches typically substitute some of the originally acquired geometric information by textures. Moreover, mesh reconstruction is a costly process that, depending on the size of the data set, can take days to compute.

Our goal was to preserve as much geometric detail of our data sets as possible and to maintain the precise measurements taken during acquisition. Since each sample of the point cloud is already associated with a position and color information, meshing the point cloud would introduce unnecessary memory overhead and processing time. In some cases, e.g. the woodland of the valley, we even found that 3D point clouds provide a more authentic visual appearance than meshed representations.

7.1 Large Multi-scale 3D Scenes

The size of the Valcamonica scanning data sums up to hundreds of gigabyte. The model contains 14 billion 3D points which can neither be processed in real-time nor stored on current graphics hardware. Moreover, rendering all points may result in massive overdraw with disturbing aliasing artifacts. Thus, the visualization requires an output-sensitive approach that prepares the data to be displayed at an appropriate geometrical resolution which approximately matches the resolution of the physical display.

Therefore, we employ state-of-the-art level-of-detail (LOD) data structures (Goswami et al., 2010, 2013) to reduce the amount of data used for rendering. In an offline pre-process, our data sets are simplified and organized hierarchically using kd-tree space partitioning data structures. At run-time, we adaptively select and update cuts through these hierarchies using a view-dependent level-of-detail selection scheme inspired by (Carmona & Froehlich, 2011). This approach facilitates the restriction of the resident data to the parts of the scene that are within the viewing frustum. Our out-of-core data management is based on a two-level cache hierarchy, designed to stream the required parts of the scene on demand from external memories (e.g. network or SSD) to the graphics memory.

7.2 Appearance-preserving Simplification

The inner nodes of our level-of-detail hierarchies contain simplified representations of the original data. When investigated at differing levels of detail, the visual appearance of the model should be preserved and never change abruptly. Therefore, we require the output of the simplification algorithm to represent the surface characteristics of the finer representations as accurately as possible.

A number of simplification strategies for point-based data have been proposed (Goswami et al., 2010; Pauly et al., 2002; Goswami et al., 2013; Alexa et al., 2001). However, we found few of them well-suited for scanned data sets as we need to maintain a uniform point distribution for the interior nodes of the LOD-hierarchies to preserve the visual appearance. In our case, only hierarchical clustering (Pauly et al., 2002) and normal deviation clustering (Goswami et al., 2013) worked well. In particular, we observed that normal deviation clustering outperformed other simplification strategies, because the employment of a regular grid tends to preserve uniform point distributions and the normal deviation criterion leads to a preservation of geometric features.

7.3 Point Cloud Rendering

In our system, the 3D points are rendered as oriented, circular shapes also referred to as surfels (Pfister et al., 2000). For visual fidelity, overlapping surfels can be blended smoothly during rendering using Gaussian weighting (Botsch et al., 2005). Although surfels are not interconnected, blended surfels result in a surface approximation if appropriate radii and normals are either known or computed. However, the estimation of appropriate surfel radii is rather intricate as the position, orientation and shape of the surfels need to be considered. If radii are overestimated, many surfels overlap unnecessarily and the number of fragments projected onto a pixel may become a bottleneck. On the other hand, an underestimation of radii may lead to holes in the surface. Therefore, radii need to be computed such that the surface appears to be watertight, but with minimum overlap between surfels. As a solution, we compute the natural neighbors for any given surfel and assign half the distance to the farthest neighbor as radius. For uniformly-sampled point clouds, this approach provides for closed surface appearance with a minimum overlap between points. If the data set contains outliers or the points are not evenly distributed over the scanned surface, we recommend to employ an intermediate supersampling of the point cloud to regularize point distribution and improve rendering performance.

8 3D INTERACTION TECHNIQUES

The visual analysis of large 3D data sets requires appropriate interaction tools and techniques. Collaborative systems based on immersive 3D visualization such as the 3D-Pitoti Scientists' Lab open a large design space for novel developments in that regard. This includes adaptations of many established analysis methods, e.g., from geographical information systems (GIS), for application in the novel environment. Our research, so far, has focused on enabling technologies and exemplary interaction techniques that demonstrate the potentials of collaborative and immersive 3D visualization. Some examples are described in the following sections

8.1 Ray-based Interaction with large 3D Point Clouds

Pointing with a virtual ray is one of the most fundamental 3D input techniques. It is generally used for the selection of objects and locations as well as for coarse object manipulation. In the 3D-Pitoti Scientists' Lab we use ray pointers for the placement of reference points in the scene, for measurements, and to define 3D surface regions for illustrative visualization of their geometric characteristics (Figure 8(c)). These techniques are based on intersection tests of a virtual ray with the displayed geometry. Also navigation techniques can build on such intersection tests, e.g. to measure the distance to the virtual ground level and maintain it during locomotion.

Our large 3D point cloud data requires specialized intersection tests to this end. The geometry, represented through surfels, is not completely watertight, and therefore small gaps and discontinuities may result in poor intersection quality. Due to the level-of-detail hierarchies and out-of-core data management for large scenes, the geometry data residing in RAM depends on the current configuration of views. The interaction techniques, however, and in particular ground following navigation, require robust intersection tests also with geometry that might be located outside of the viewing frustum.

We exploit the level-of-detail data hierarchy to accelerate intersections. Only the finest nodes whose associated geometry information is present in the main memory are considered for intersection tests with the contained surfels. All planes defined by the set of surfels are intersected and the result closest to any center of these surfels is chosen. This ensures robust intersections, despite small holes in the surface representation. If the available level of detail is too coarse, because the intersection area is out of view, we resort to the boundary volumes of our level-of-detail hierarchy which always resides in main memory. In these cases, we traverse the hierarchies down to the leaf level in order to obtain more accurate intersection results.

Point clouds of 3D-scanned surfaces are inherently noisy. Therefore, the above described intersection method results in depth deviations between intersection tests at directly adjacent locations – even in case of an originally flat surface. For some techniques, such as ground-following navigation, the average intersection value of multiple rays thus offers a better approximation of the actual surface. Our bundle pick technique, therefore, performs intersection tests with 8 to 16 rays in parallel and fits a plane through the resulting 3D points.

8.2 Multi-scale Group Navigation

Group navigation is a predominant example for the specific interaction design requirements of collaborative systems. The group wants to navigate together, but all participants may have different ideas where to go next. Therefore, all users must have access to provide group navigation input. User actions in this regard must be clearly visible and foster negotiation, even interference, if necessary. The group navigation input must be applied smoothly and the resulting change of location must be comprehensible for everybody involved.

8.2.1 Egocentric Group Navigation

Our large vertical screen, affords an immersive experience of the displayed content from an egocentric perspective. When navigating through the scene, all participants are moved together in order to maintain workspace coherence. The virtual locomotion is controlled with a centrally accessible group navigation device. It offers seven degrees of freedom (DOF) for 3D rotation, translation and uniform scaling. The physical presence of the large device and the operating user fosters awareness (Figure 5). Its operation can be combined with other input actions to facilitate group navigation. We use a 3D pointer, for example, to redefine the center of view rotation and scaling. Per default we scale and rotate around the large 3D trackball of the group navigation-device, but this transformation center can be set anywhere in the virtual environment (Figure 5(a)).

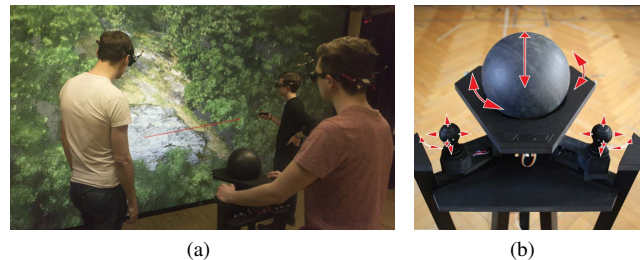


Figure 5: A centrally placed group input-device allows to navigate the shared interaction space through the virtual environment with seven degrees of freedom (3D rotation, 3D translation, and uniform scaling). While one user operates the device, another one can define suitable reference points for rotation and scaling with a ray pointer (a). The device is equipped with two 3D joysticks and a 3D trackball (b) Each joystick enables 3D translation and head rotation. In combination they support uniform scaling too.

Such scaling around subsequently specified reference locations is an efficient method for multiscale navigation in large virtual environments. Decreasing the size of the environment provides overview for the selection of the next target location. The miniaturization also brings it into reach, where it remains during magnification. Accurate view positioning requires only small maneuvering movements thereafter. Collaborators usually distribute the operations of navigation and pointing input among two participants.

We also implemented automatic scale adjustments during virtual navigation similar to earlier proposed techniques for stereoscopic 3D environments (e.g. Carvalho et al., 2011; Argelaguet & Maignant, 2016; Cho et al., 2017). More specifically, we use a cubemap to encode distances between the navigating user and the environment (McCrae et al., 2009) and adjust motion velocity as well as scaling in such a way that the surrounding environment remains in a reasonable depth range for accurate and unimpaired stereo perception². This allows users to focus on translation and rotation input for steering through the virtual environment, while an appropriate scale is automatically specified. Moving closer to a specific detail, increases its size, while moving away improves overview through miniaturization.

The technique works well for single users when the scaling center is set to the user's head position. Scaling the environment around the center between both eyes is the inverse operation to changing the virtual eye separation. In case of group navigation, however, there is no ideal transformation center for automatic scaling. It can be set between both eyes of the operating user or to the group navigation device centrally in front of the shared display. However,

²The depth range depends on the type and size of display.

the further a user is away from this reference position the more irritating is the resulting motion flow of scaling. We found that group navigation through a landscape environment such as virtual Valcamonica is more comfortable and effective through explicit scaling around subsequently specified target locations as described above even though it is more laborious.

8.2.2 Allocentric Scene Exploration

On the tabletop display we offer the same 7-DOF navigation functionality through multi-touch input. Compared to the controller-based navigation on the large powerwall, the touch table affords different views and exploration paths. Changing the view towards the displayed 3D content on the tabletop follows the metaphor of object manipulation. The top-down view affirms the perception of moving a miniature environment with one's hands. Consequently, 3D navigation on the tabletop does not induce the visual perception of self motion which would be in conflict with vestibular cues of standing still in front of the tabletop (as known from more immersive displays). We observed that most users prefer the table to search for interesting locations in the 3D map visualization of virtual Valcamonica.

This preference may also be related to the intuitive and approachable multitouch navigation. Our implementation largely corresponds to touch gestures for 2D rotation, scale, and translation (RST) as known from mobile touch devices. The main difference is that scaling can only be performed with a two-handed pinch gesture (Figure 6(b)). We constrained this functionality to facilitate data exploration at constant scale. One-handed input can only induce planar rotation and translation. Following the touch recognition approach of Ewerling et al. (Ewerling et al., 2012), we associate all touch inputs on our tabletop device with the corresponding hands.

Another benefit of multitouch input for multiscale 3D navigation is that the center of rotation and scaling is implicitly defined and does not require additional input. Similar to De la Rivière et al. (2010), we select reference positions on the 3D model through intersection with rays that extend through the centers of the operating hands in orthogonal direction to the touched display. The closest intersection point to the display is chosen as the transformation center (Figure 6).

The established RST gesture set supports only four degrees of freedom while seven need to be controlled for multiscale 3D navigation. We therefore introduced additional gestures for adjustments of height and full 3D rotation. Earlier suggested gestures for vertical translation generally correspond to the pinch-zoom gesture which we apply for scaling (Hancock et al., 2009; Strothoff et al., 2011). Similarly, the tilting part of suggested 3D rotation gestures closely resembles 2D translation input (Martinet et al., 2010; Herrlich et al., 2011; Liu et al., 2012; Cohé & Hachet, 2012). With a series of user studies, we identified subtle differences that support automatic differentiation between these different types of input. In short, we analyze movement symmetry to distinguish scaling (symmetric) from vertical translation (asymmetric) and we consider the number of involved fingers to distinguish planar transformations with one hand from 3D rotation input with a two-finger precision grasp (Figure 6).

8.3 Photoportals

Our system presents 3D scans of prehistoric petroglyphs embedded in a model of the surrounding environment. This is an important feature for archaeological analysis and interpretation, since the actual artworks are immovably bound to the places of their creation. Movable records such as photos and tracings facilitate the detailed analysis and the comparison of individual figures, but show the captured art in isolation from its spatial context.

In the virtual model of Valcamonica, individual Pitoti could be separated from the surrounding and moved as desired, but this would contradict the requirement to view the art in context. Instead, we implemented the Photoportals technique of Kunert et al. (2014). This technique follows the metaphor of photography to create and exchange specific views, but, the virtual Photoportals are dynamic 3D views that can always be adapted and that support different perspectives, e.g., at the surrounding of a captured detail (Figure 7).

Sharing perspectives can be highly relevant to achieve mutual understanding, in particular if certain objects and interrelations can only be seen from specific viewpoints. Photoportals allow users to capture the respective views and immediately show them to others. They implicitly also serve as a cutting plane to reveal occluded geometry and create cross-section views (Figure 7(b)). This facilitates group discussion of areas and views that cannot be reached comfortably with group navigation techniques.

Photoportals establish direct spatial links to different locations in the scene, that serve as visual references, but also, they can be used for navigation. We implemented two techniques to enter selected locations. Individual users can simply put their head through the Photoportal to better perceive the scene on the other side of the virtual window (Figure 7(c)). Teleporting the whole group to this location means to replace the view of the shared physical display with that of the Photoportal. Correspondingly, we increase the size of the virtual window until it covers the whole screen. This procedure can be applied to both physical 3D displays individually.

8.4 Virtual Illumination and Illustrative Visualization

The Pitoti figures can be best seen with incident light from a shallow angle as it occurs during the "magic hours" of dawn and sunset. The shadows in the deeper regions of the petroglyphs increase the visual contrast and their movement seems to animate the figures. The archaeologists reported that they use artificial light sources to achieve the best possible appearance of these artworks in Valcamonica (Figure 3(c)). In the 3D-Pitoti Scientists' Lab we implemented a virtual torch that allows the same visual effects in virtual reality with the additional possibility to adapt visual material properties (Figure 8(b)). This functionality, similar as described by Di Franco et al. (2015), in combination with the possibility to remove texture and increase the perceived size of the figures allows a much better analysis of their geometric structure. The virtual torch works across all viewing contexts, i.e., the powerwall, the tabletop and the Photoportals.

Computer supported visualization offers many more options to emphasize the subtle artworks pecked into the rock surfaces (Lawonn et al., 2017). We implemented exemplary visualization techniques for color-coded illustrations of relative surface depth and its gradient. These can be applied to selected surface areas to facilitate comparisons between peck marks (Figure 8(c)). Moreover, we use color overlays to highlight regions that have been manually or automatically classified as anthropogenic traces (Zepelzauer et al., 2015) (Figure 8(d)).

8.5 Multi-Window 3D Interaction

The availability of multiple physical and virtual 3D viewing windows allows users to simultaneously explore different aspects of a 3D scene and compare them. Navigating through the virtual environment with one display does not affect the view provided by others, hence subgroups of users can explore different aspects of the scene separately. The efficiency and fluency of group interaction can benefit from such temporary phases of divergence. Users can explore the same data from different perspectives and later present their individual insights to others.

Despite the different views, users expect seamless 3D interaction across all collocated displays (Figure 9). This requires concurrent representations of users and interaction devices for each separate

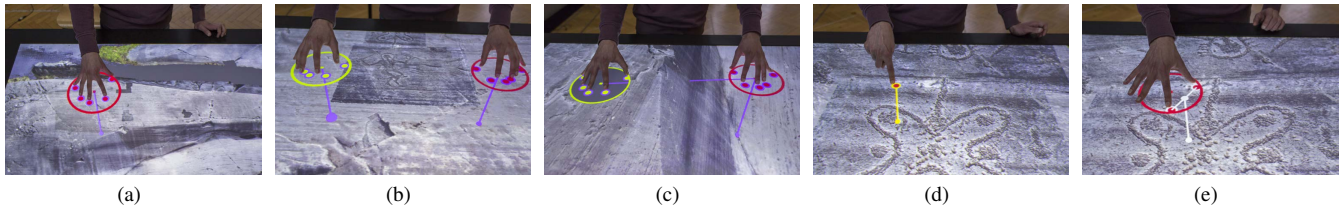


Figure 6: The tabletop enables 3D multiscale navigation with multitouch input. Two-handed input gestures on the display surface directly map to 2D transformations (a&b). In our case, scaling is only available with two-handed input (b). The pinch between both hands can also operate vertical translation, if applied asymmetrically (c). We also take into account the number of fingers per hand: one finger does not change the scene, but it controls a virtual ray to indicate features behind the screen (d), one-handed input with two spread fingers operates full 3D rotation (e), and one hand with more than two fingers operates translation (a).

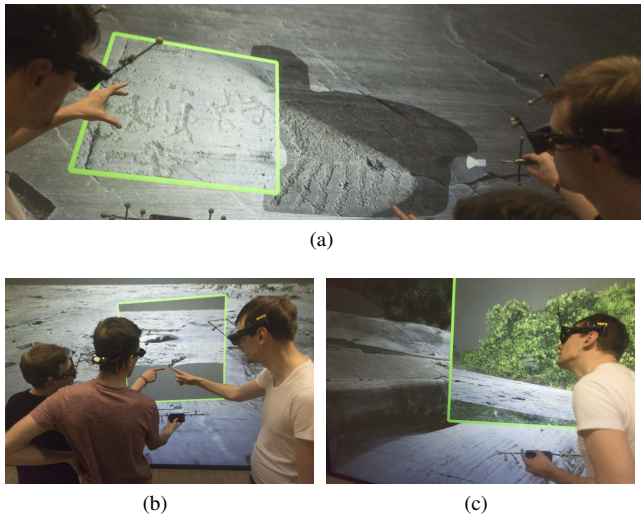


Figure 7: Photoportals build on the metaphor of photography to create additional virtual viewports. They can be used, for example, to compare views of distant locations (a), to create cross section views (b), or as a navigation window to remote place in the virtual environment (c).

display. Any user input in the physical workspace must be considered for interaction with the scene sections shown on each of the involved displays.

We developed a multi-context 3D interaction framework that implements and manages these concurrent user inputs based on tool-specific heuristics for the selection of active contexts. Photoportals, for example, must be visible on all collocated displays simultaneously, while their view capturing and teleportation functionality can only apply to a single context at a time. This allows users to capture a view on one display and apply it to another (Figure 9(a)). The effect of a handheld virtual light source, instead, should be applied to all contexts in reach, to enable the comparison of distant Pitoti figures under similar lighting conditions (Figure 9(b)).

9 EVALUATION

We evaluated the technical performance of our rendering system as well as the usability of the end-to-end visualization system. The performance and visual quality of the real-time rendering technology was quantified for an exemplary multiscale 3D navigation path. The overall usefulness and usability of the 3D-Pitoti Scientists Lab was evaluated qualitatively by five domain experts using the system for one or two days.

9.1 Rendering Performance and Discussion

The evaluation of rendering performance and quality was conducted on an *Intel Xeon* CPU with 6 cores at 3.5 GHz, 128 GB main memory along with a *Nvidia Titan X* graphics card with 12 GB video memory and a screen resolution of 1920x1080. Our data set of the Valcamonica valley covers a real-world extent of more than 1.5 kilometers and has a total size of more than 400 GB. The model consists of about 14 billion points – the large-scale scans of the valley amount to 10 billion points while the detailed scans of rocks and rock-art amount to 4 billion points.

Figure 10 shows a performance test of our rendering system. The measurements were taken while navigating from a medium-range overview of the valley to a close-up view of a single rock-art engraving. During this navigation, the draw time varies between 10 ms for the overview and about 40 ms for the close-up. For all views, the performance allows for an interactive exploration of the model. The chosen level-of-detail representation used for rendering varies between 2 and 8 million points and is only a small fraction compared to the total number of 14 billion points. For the valley view, fewer points are drawn because of the comparably low scanning resolution. In general, the number of rendered points remains quite constant and mostly depends on the screen resolution because of the output-sensitive level-of-detail selection.

In general, a naïve implementation of stereoscopic rendering increases the draw times by a factor of two since only the cut update of the multi-resolution representation can be shared between the left and right eye view. However, there are various approaches that achieve similar frame rates also for stereoscopic rendering. First, parallel rendering of both eyes can be achieved by using one dedicated graphics device for each eye and each individual user. In this case, the stereoscopic display of the resulting image pairs requires a synchronized frame swap. Secondly, preliminary tests have shown that latest image-warping techniques (Schollmeyer et al., 2017) can be used to produce stereoscopic views with only little overhead. For static 3D scans, the potential visual artefacts remain unnoticeable by most users. Furthermore, the latest graphics hardware is capable of reducing the overhead of geometry processing by generating stereo views based on horizontal disparities. This works well for head-mounted displays, however, for projection-based virtual reality, its use is rather limited because asymmetric view frusta and arbitrary head rotations are not accounted for.

Finally, we evaluated the image error introduced by using a level-of-detail representation. Image differences were measured using the structural similarity (SSIM) method (Wang et al., 2004) because its results match better with the human visual perception. Figure 11 shows an image comparison between a level-of-detail rendering and ground truth. For generating the ground-truth image, the highest available detail was rendered in high resolution in order to prevent aliasing artifacts and then downsampled to match the resolution used for comparison. The results show that the level-

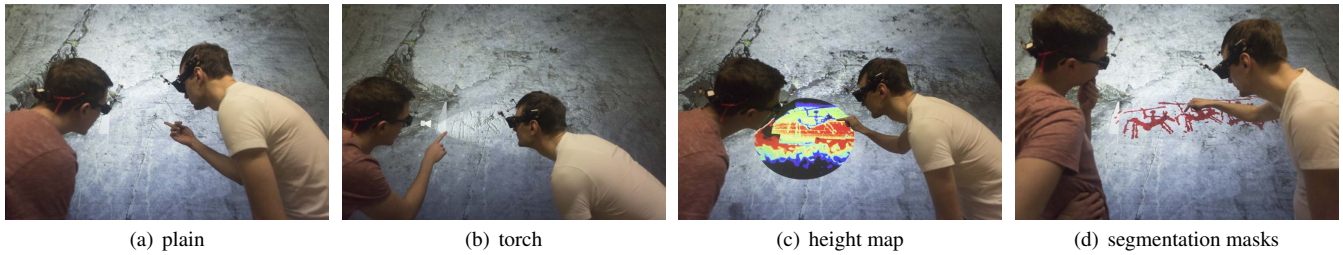


Figure 8: Two users are discussing the differences between natural and anthropogenic traces on the rock surface. They use pointing and tracing gestures to indicate, the structures they are referring to. Some of the petroglyphs are hardly visible in the plain 3D model (a). Virtual illumination emphasizes the pecked 3D structures (b). The surface geometry can be further emphasized with illustrative visualization techniques, e.g. a height map (c). If segmentation masks are available, they can be used to highlight the artwork (d).

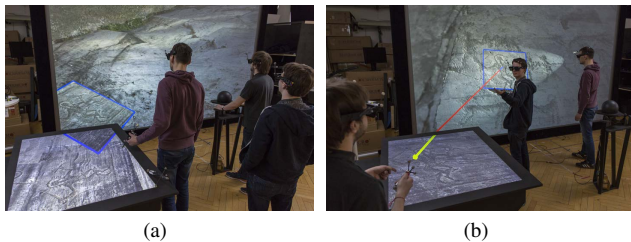


Figure 9: The 3D-Pitoti Scientists' Lab offers multiple independent 3D viewing windows, e.g., to explore different locations and scales of the same scene in parallel. We create a coherent 3D interaction space between these different virtual interaction contexts. This allows, for example, to capture views at one physical display and move it over to another one using a Photoportal (a). Also pointing rays and virtual light sources consistently appear and are applied across multiple displays and contexts (b).

of-detail rendering often reaches a very high similarity with only about 10%-25% of the geometry costs compared to the ground truth views. For the views of the rock (Figure 11(d)) and a close-up of a Pitoti (Figure 11(a)), the high image similarities (SSIM=0.9823, SSIM=0.9877) indicate that the difference is hardly noticeable by the human eye. The lower spatial resolution of the valley scan (Figure 11(g)) leads to a moderate amount of visual artefacts caused by the simplification algorithm used to build the level-of-detail hierarchies. However, these artefacts are hardly avoidable due to the high frequencies in the forest areas.

9.2 Expert Review

We invited five archaeologists (25-64 years, two female) to evaluate our collaborative 3D visualization lab. Two of them were part of the project and also involved in the requirements analysis. The other three experienced the system for the first time. All five participants had extensive knowledge of rock art. The experiments were organized in two sessions, one with two participants lasting a single day and another one with three participants lasting two days. The participants were invited based on their individual interest in using the technology and were not paid for taking part in the study.

We started both sessions with a hands-on introduction to the system and its functionalities. It took about half an hour per participant to operate each feature of the system at least once. Thereafter we discussed the specific archaeological interests of each participant and how the system could be used in this context. We agreed to explore corresponding research questions in an unstructured collaboration and finally capture a presentation of the findings on video (using a video camera with a dedicated 3D-tracked view of the dis-

played 3D scene). As a result, the participants were motivated to use the system for their own benefits and evaluate its usefulness on this basis. Extracts from these videos have been submitted with this paper.

During breaks between the operation of the system we conducted unstructured group interviews about the relevance and usability of particular system features. We asked the participants to take home a questionnaire regarding the usefulness and usability of the system including a system usability scale (SUS) (Brooke, 1996) and return their responses on the next day after they had some time of reflection over their experience with the system. In the following sections we report on the results concerning only the particular functionalities of our visualization system

9.2.1 Interview Results

During the interviews all participants strongly agreed on the general usefulness of interactive immersive 3D visualizations as provided by the system. When asked how they would use such system in their own research, they expressed its value for presenting the data and exploring it with a very realistic appearance at any time and any place. The specific usefulness for the geometrical analysis, of superimpositions of multiple petroglyphs for example, was mentioned. Considering that such analysis can also be accomplished with alternative 2D visualization techniques such as section views and depth maps, the participants emphasized the advantage of seeing such additional information directly applied to the actual 3D geometry, which supports cross validation of the information. One participant highlighted the benefits of experiencing the scanned sites and the peck marks in relation to her own body. The immersive 3D environment enables the exploration from perspectives that resemble the process of approaching the physical location.

The archaeologists agreed that collaboration support was one of the fundamentally relevant features of the system, since it facilitated mutual exchange about opinions and perceived geometrical characteristics. One participant mentioned that this feature would be even more powerful if the involved collaborators could also be located remotely and meet via telepresence technologies at the scanned site.

The tabletop was considered more appropriate for collaborative analysis and discussions, while the powerwall was considered useful to gather individual impressions of the environment or give presentations to others. Standing around the tabletop device facilitated face-to-face communication about the data and the multitouch input was immediately and equally accessible to all involved users. This observation is in line with earlier research by Rogers and Lindley, who found that shared horizontal displays can support higher awareness and more frequent changes of roles than vertical ones (Rogers & Lindley, 2004). The powerwall allowed only for one controlling user at a time who had to be standing at the group navigation

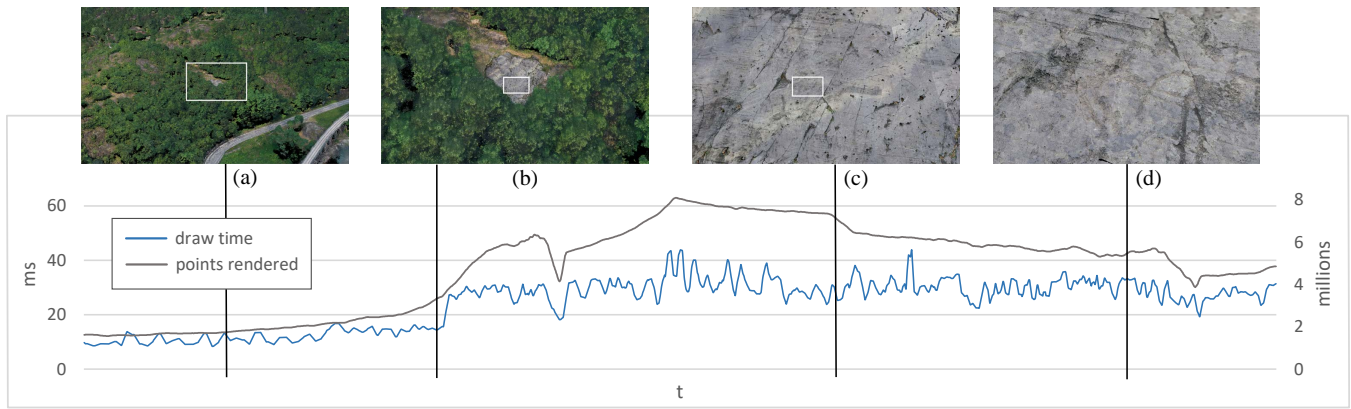


Figure 10: This Figure shows the performance of our rendering system for a medium- to close-range navigation. The vertical lines correspond to the views shown above. Starting from an overview of the rock-art site Seradina Rock 12C (a), we navigate towards the rock panel (b, c) and stop with a close-up of a single Pitoti(d). The draw time (blue) varies between 10 ms to 40 ms depending on the current view and the number of points (shown as gray line) in the corresponding level-of-detail representation.

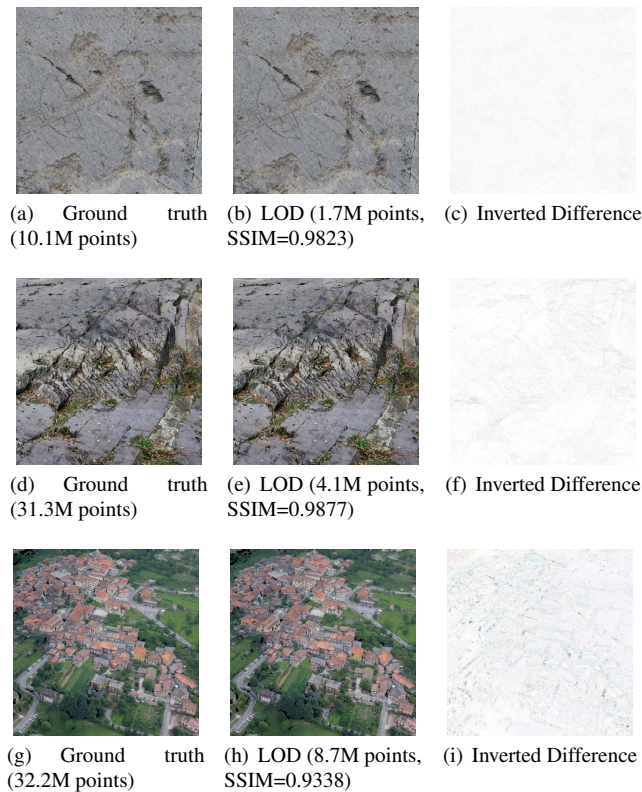


Figure 11: These images show the result of our output-sensitive rendering compared to ground truth along with the inverted image differences. For all views, the LOD representation contains considerably fewer points, but in most cases the image difference is almost imperceptible. The image error is higher for the valley ((g)- (i)) because the simplification algorithm used for LOD generation does not deal well with high-frequency and noisy scanning data.

device. Thus, the handover of navigation control was also easier and faster with the tabletop.

Our domain experts appreciated the virtual 3D viewing windows (Photoportals). They found this feature helpful and easy to use, in particular to capture, adapt and exchange secondary views for the comparison of different figures in different locations. Photoportals were also used to create section views, although the participants reported that this usage would benefit from higher manipulation accuracy. The unconstrained 3D placement of cutting planes was challenging. They also asked for the possibility to organize Photoportals like virtual post-it notes on the projection screens. The functionality to directly enter a captured location was considered easy to understand and use, however, our implementation did not require explicit group agreement which led to a few unexpected location changes for others.

When asked about the multi-context functionality, e.g., using the virtual flashlight across displays, our experts agreed that this corresponded to their expectation. They had difficulties to imagine a different behavior, i.e. that tools would change their mode or state or affect the scene only in one of the 3D views. They also appreciated the portability of virtual displays between physical ones and requested extensions to their own mobile devices in order to take captured views from an analysis session home as 2D snapshots.

9.2.2 Usability ratings

The mean score obtained from the system usability scale was 85 with a standard deviation of 7.42. This result, clearly above the average value of 68, reflects that all our testers learned to operate the complex multi-display 3D interaction system³.

The archaeological experts had received an introduction to the technology before arrival at our lab using text, photos and also video clips. We provided this information to trigger their interest. Therefore, our questionnaire also asked the participants to compare the actual experience with their initial expectations. On a scale from 1 (much worse than I expected) to 7 (much better than I expected), three participants rated 6, one user rated 5 and one rated his experience as 7. They explained their positive responses with the effective integration of the tabletop display and the powerwall, as well as the different tools in general. Also the system responsiveness and its support for collaboration were highly appreciated in the comments.

Further positive factors included the large range of interaction possibilities. One user said: “As an experiential set of devices, it is

³<http://www.measuringu.com/sus.php>

fantastic. There is simply a strong element of 'enjoyability' in vicariously experiencing the landscape and art in a 3D environment. The interface also promotes group engagement which trumps other systems such as the Oculus Rift in that respect." Their comments also indicated that the very positive evaluation did not only reflect their experience with our test system, but that it was influenced by their expectations of a potential future product based on our developments.

10 DISCUSSION

3D scanning and interactive 3D data visualization have become established methods for data recording and exchange in archaeology (Chapman et al., 2013). In this project we realized a novel approach for the interactive collaborative exploration of 3D scanned rock art in a coherent multi-scale virtual environment. An expert review of our experimental implementation approved it as a viable extension of traditional rock-art research methods. While the virtual environment cannot replace the original place with its atmosphere and aura, it offers other benefits for archaeological research, including; better accessibility, controlled environment variables, and easier exchange with other researchers.

The experimental system was co-designed with archaeologists and many of their requested functionalities were implemented, e.g., to support controlled lighting, magnification, and augmented data visualization. We also integrated early results of computer supported data interpretation with machine learning techniques (Zeppelzauer et al., 2015). The developed methods for visual analysis mimic those used by archaeologists in the field, but under controlled conditions. For example, Photoportals can be used like a magnification lens. The virtual flashlight allows the immediate exploration of dynamic lighting effects on the rock art. Similar features have been implemented in other virtual representations of cultural heritage to permit an enhanced or perception of surface qualities (Di Franco et al., 2015). In our case, the dynamic appearance of the illuminated rock art is considered an essential feature of the art and thus forms part of the interpretation of rock art scenes (Chippindale & Baker, 2012).

A particular feature well received by our participants is that all actions taken whilst exploring the virtual environment, including the visual effects and perceptions can be recorded and fully reproduced at any time. This matters for 3D digital cultural heritage applications as it has been noted that "There is no specialized software for CH professionals to document their work in 3D." (Soler et al., 2017, p.49), and our system offers some potential in this respect. However we did not implement all features in our demonstrator system that research scientists considered would be needed. For example, our archaeologists also requested specific functionalities such as a spatiotemporally correct simulation of sunlight in the valley, or options to export the extracted information to other systems for analysis and presentation. This highlights the need to support functional extensions and interfaces to other sources of data and processing methods, e.g., geographical information systems (GIS), 3D modeling and animation software, and statistics software.

During an extensive evaluation, our rock-art experts expressed their appreciation for the possibilities and showed how they could use the system to convey their knowledge and their interpretations of the prehistoric art to others. Our review panel also reported that visualization features offered by the 3D-Pitoti Scientists' Lab, could add great value to their research since they allow them to interrogate the data in ways that are not possible or easy using traditional methods. One example is magnification, which permits new types of enquiry regarding the artwork. High resolution magnified visualization reveals details of the way in which the figure was created: individual peck marks can be measured and information gleaned about the tool used and force applied to each peck. Potentially, this could allow archaeologists to identify the pecking styles

of individual artists. Photoportals enable further for comparison of artworks located at different sites across the valley as they allow the researcher to 'carry' the petroglyphs around the virtual valley, while maintaining their individual spatial context.

The system presented and tested so far builds on custom technologies that are not readily available to the archaeological research community. The integration of the presented collaborative visualization and interaction capabilities with existing ecosystems of analytical software and research databases (e.g. Soler et al., 2017) is a prime topic for future developments. As an alternative to the large multi-user 3D projection systems, VR-headsets should be considered. However, we believe that this approach complicates the collaborative data exploration as a head-mounted display obscures the perception of one's own body and the presence of others, which in turn affects visual perception, collaboration and mutual awareness. Real-time 3D scanning technologies as developed for telepresence applications could minimize this shortcoming (Maimone & Fuchs, 2011; Beck et al., 2013; Dou et al., 2016).

11 CONCLUSION AND FUTURE WORK

We presented a novel end-to-end system for the collaborative visual analysis of extremely large 3D scanning data in multi-user virtual reality. Our developments contribute conceptual and technical solutions to the involved challenges of appearance-preserving real-time rendering and collaborative 3D interaction. Our technologies were successfully integrated in a prototypical visualization infrastructure for collaborative 3D data analysis. A technical evaluation and a usability study with domain experts confirmed expected benefits of the system components and the system as a whole.

We are currently working on further improvements of our technologies, e.g., to support faster rendering at higher image resolution, but also to extend the range of interactive functionalities for visual analysis. We aim to further leverage the possibilities of collaborative interaction in that regard. We are also planning to integrate our novel facilities for visual 3D data analysis with established research methods and workflows in the cultural heritage domain. The very next step is the adaptation of our systems to support the exploration of other large-scale cultural heritage data like reconstructed historical architecture, cities and landscapes. Several concrete challenges towards these goals can already be estimated.

Faster image generation for higher-resolution displays and potentially more complex data requires more efficient rendering techniques. We will consider the triangulation of the scanning data where appropriate, but we are also working on improved techniques to omit the rendering of occluded parts of the scene.

On the side of the user interface and collaborative interaction techniques, we are currently working on a sensible integration of head-mounted displays that are a more readily available alternative to multi-user 3D projection technology. However, we believe that meaningful 3D data exploration and effective collaboration builds on the veridical perception of others and one's own body in relation to the virtual environment. Using head-mounted displays, therefore, requires rich 3D avatar representations of all involved users.

This relates to current work towards remote collaboration. The foundational technologies for group-to-group 3D telepresence are available (Maimone & Fuchs, 2011; Beck et al., 2013; Fairchild et al., 2016; Dou et al., 2016). However, simultaneous real-time rendering of 3D video avatars and extremely large 3D scanning data at high visual quality remains challenging. Moreover, novel interface developments are necessary to support workspace awareness and the coordination of collaborative work with remote participants in large scale virtual environments.

We believe that the collaborative visual analysis of 3D scanning data in a social virtual reality has a huge potential to advance research and education methods in archaeology and related academic disciplines.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the EC FP7 project 3D-PITOTI (ICT-2011-600545). Our work was also supported by the German Federal Ministry of Education and Research (BMBF) under grant 031PT704X (project Big Data Analytics). We thank the participants of our workshops and studies, the members of the 3D-Pitoti project consortium, and the members and students of the Virtual Reality and Visualization Research group at Bauhaus-Universität Weimar (<http://www.uni-weimar.de/medien/vr>). This paper is dedicated to our dear colleague and friend Sally Shalloe who sadly died during the 3D-Pitoti project. We are indebted to Sally for her significant contribution to this work.).

REFERENCES

- Alexa, M., Behr, J., Cohen-Or, D., Fleishman, S., Levin, D., & Silva, C. T. (2001). Point set surfaces. In *Proceedings of the conference on visualization '01* (pp. 21–28). Washington, DC, USA: IEEE Computer Society. Retrieved from <http://dl.acm.org/citation.cfm?id=601671.601673>
- Alexander, C. (2011). *Valley of pitoti: Gis-based socio-spatial analysis of rock-art in valcamonica (bs), lombardy, italy* (Unpublished doctoral dissertation). Ph. D. thesis. University of Cambridge: Cambridge.
- Alexander, C., Pinz, A., & Reinbacher, C. (2015). Multi-scale 3d rock-art recording. *Digital Applications in Archaeology and Cultural Heritage*, 2(2–3), 181 - 195. Retrieved from <http://www.sciencedirect.com/science/article/pii/S2212054815000077> (Digital imaging techniques for the study of prehistoric rock art) doi: <https://doi.org/10.1016/j.daach.2015.02.003>
- Anati, E. (1961). Camonica valley: a depiction of village life in the alps from neolithic times to the birth of christ as revealed by thousands of newly found rock carvings.
- Anati, E. (1976). *Evolution and style in camunian rock art: An inquiry into the formation of european civilization* (Vol. 6). Edizioni del centro.
- Anati, E. (1977). Methods of recording and analysing rock engravings. *Studi Camuni Brescia*, 7, 3–61.
- Anati, E., & Cittadini, T. (1994). Valcamonica rock art: a new history for europe.
- Argelaguet, F., & Maignant, M. (2016). Giant: Stereoscopic-compliant multi-scale navigation in ves. In *Proceedings of the 22nd acm conference on virtual reality software and technology* (pp. 269–277). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/2993369.2993391> doi: 10.1145/2993369.2993391
- Baker, K., Greenberg, S., & Gutwin, C. (2002). Empirical development of a heuristic evaluation methodology for shared workspace groupware. In *Proceedings of the 2002 acm conference on computer supported cooperative work* (pp. 96–105). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/587078.587093> doi: 10.1145/587078.587093
- Barreau, J.-B., Nouviale, F., Gagne, R., Bernard, Y., Llinares, S., & Gouranton, V. (2015). An immersive virtual sailing on the 18th-century ship le boullongne. *Presence: Teleoperators and Virtual Environments*, 24(3), 201–219.
- Beck, S., Kunert, A., Kulik, A., & Froehlich, B. (2013, April). Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 19(4), 616–625. Retrieved from <http://dx.doi.org/10.1109/TVCG.2013.33> doi: 10.1109/TVCG.2013.33
- Benford, S., Bederson, B. B., Åkesson, K.-P., Bayon, V., Druin, A., Hansson, P., ... Taxén, G. (2000). Designing storytelling technologies to encouraging collaboration between young children. In *Proc. of chi 2000* (pp. 556–563). ACM Press. Retrieved from <http://doi.acm.org/10.1145/332040.332502> doi: 10.1145/332040.332502
- Botsch, M., Hornung, A., Zwicker, M., & Kobbelt, L. (2005). High-quality surface splatting on today's gpus. In M. Alexa, S. Rusinkiewicz, M. Pauly, & M. Zwicker (Eds.), *Spbg* (p. 17-24). Eurographics Association. Retrieved from <http://dblp.uni-trier.de/db/conf/spbg/spbg2005.html#BotschHZK05>
- Bradley, D., Boubekur, T., & Heidrich, W. (2008). Accurate multi-view reconstruction using robust binocular stereo and surface meshing. In *Computer vision and pattern recognition, 2008. cvpr 2008. ieee conference on* (pp. 1–8).
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Weerdmeester, A. Thomas, & I. L. McLelland (Eds.), *Usability evaluation in industry*. London: Taylor and Francis.
- Bruno, F., Bruno, S., De Sensi, G., Luchi, M.-L., Mancuso, S., & Muzzupappa, M. (2010). From 3d reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*, 11(1), 42–49.
- Carmona, R., & Froehlich, B. (2011). Error-controlled real-time cut updates for multi-resolution volume rendering. *Computers & Graphics*, 35(4), 931-944. Retrieved from <http://dblp.uni-trier.de/db/journals/cg/cg35.html#CarmonaF11>
- Carvalho, F., Trindade, D. R., Dam, P. F., Raposo, A., & d. Santos, I. H. F. (2011, May). Dynamic adjustment of stereo parameters for virtual reality tools. In *2011 xiii symposium on virtual reality* (p. 66-72). doi: 10.1109/SVR.2011.30
- Chapman, H., Baldwin, E., Moulden, H., & Lobb, M. (2013). More than just a sum of the points: Re-thinking the value of laser scanning data. In E. Ch'ng, V. Gaffney, & H. Chapman (Eds.), *Visual heritage in the digital age* (pp. 15–31). London: Springer London. Retrieved from https://doi.org/10.1007/978-1-4471-5535-5_2 doi: 10.1007/978-1-4471-5535-5_2
- Chase, A. F., Chase, D. Z., Weishampel, J. F., Drake, J. B., Shrestha, R. L., Slatton, K. C., ... Carter, W. E. (2011). Airborne lidar, archaeology, and the ancient maya landscape at caracol, belize. *Journal of Archaeological Science*, 38(2), 387 - 398. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305440310003286> doi: <https://doi.org/10.1016/j.jas.2010.09.018>
- Chippindale, C., & Baker, F. (2012). *Pitoti: digital rock-art from prehistoric europe: heritage, film, archaeology*. Skira.
- Ch'ng, E., & Gaffney, V. L. (2013). Simulation and visualisation of agent survival and settlement behaviours in the hunter-gatherer colonisation of mesolithic landscapes. In E. Ch'ng, V. Gaffney, & H. Chapman (Eds.), *Visual heritage in the digital age* (pp. 235–258). London: Springer London. Retrieved

- from https://doi.org/10.1007/978-1-4471-5535-5_12 doi: 10.1007/978-1-4471-5535-5_12
- Cho, I., Li, J., & Wartell, Z. (2017). Multi-scale 7dof view adjustment. *IEEE Transactions on Visualization and Computer Graphics, PP(99)*, 1-1. doi: 10.1109/TVCG.2017.2668405
- Cohé, A., & Hachet, M. (2012). Beyond the mouse: Understanding user gestures for manipulating 3d objects from touchscreen inputs. *Computers & Graphics, 36(8)*, 1119 - 1131. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0097849312001562> (Graphics Interaction Virtual Environments and Applications 2012) doi: <http://dx.doi.org/10.1016/j.cag.2012.09.004>
- De la Rivière, J. B., Dittlo, N., Orvain, E., Kervégant, C., Courtois, M., & Da Luz, T. (2010). ilight 3d touch: A multi-view multitouch surface for 3d content visualization and viewpoint sharing. In *Acm international conference on interactive tabletops and surfaces* (pp. 312–312). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/1936652.1936740> doi: 10.1145/1936652.1936740
- Di Franco, P. D. G., Camporesi, C., Galeazzi, F., & Kallmann, M. (2015). 3d printing and immersive visualization for improved perception of ancient artifacts. *Presence: Teleoperators and Virtual Environments, 24(3)*, 243-264. Retrieved from https://doi.org/10.1162/PRES_a_00229 doi: 10.1162/PRES_a_00229
- Domingo, I., Villaverde, V., López-Montalvo, E., Lerma, J. L., & Cabrelles, M. (2013). Latest developments in rock art recording: towards an integral documentation of levantine rock art sites combining 2d and 3d recording techniques. *Journal of Archaeological Science, 40(4)*, 1879 - 1889. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305440312005195> doi: <https://doi.org/10.1016/j.jas.2012.11.024>
- Dou, M., Khamis, S., Degtyarev, Y., Davidson, P., Fanello, S. R., Kowdle, A., ... Izadi, S. (2016, July). Fusion4d: Real-time performance capture of challenging scenes. *ACM Trans. Graph., 35(4)*, 114:1–114:13. Retrieved from <http://doi.acm.org/10.1145/2897824.2925969> doi: 10.1145/2897824.2925969
- Ewerling, P., Kulik, A., & Fröhlich, B. (2012). Finger and hand detection for multi-touch interfaces based on maximally stable extremal regions. In *Proceedings of the 2012 acm international conference on interactive tabletops and surfaces* (pp. 173–182). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/2396636.2396663> doi: 10.1145/2396636.2396663
- Fairchild, A. J., Campion, S. P., García, A. S., Wolff, R., Fernando, T., & Roberts, D. J. (2016). A mixed reality telepresence system for collaborative space operation. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Fossati, A. (1991). L'età del ferro nelle incisioni rupestri della valcamonica. *Immagini di una aristocrazia dell'età del Ferro nell'arte rupestre camuna*, 11–71.
- Fossati, A. E. (2008). Following arianna's thread: Symbolic figures at female rock art sites at naquane and in valle, valcamonica, italy.
- Goswami, P., Erol, F., Mukhi, R., Pajarola, R., & Gobbetti, E. (2013). An efficient multi-resolution framework for high quality interactive rendering of massive point clouds using multi-way kd-trees. *The Visual Computer, 29(1)*, 69–83.
- Goswami, P., Zhang, Y., Pajarola, R., & Gobbetti, E. (2010). High quality interactive rendering of massive point models using multi-way kd-trees. In *Computer graphics and applications (pg), 2010 18th pacific conference on* (pp. 93–100).
- Grosman, L., Smikt, O., & Smilansky, U. (2008). On the application of 3-d scanning technology for the documentation and typology of lithic artifacts. *Journal of Archaeological Science, 35(12)*, 3101 - 3110. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305440308001398> doi: <https://doi.org/10.1016/j.jas.2008.06.011>
- Gutwin, C., & Greenberg, S. (2002, November). A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work, 11(3)*, 411–446. Retrieved from <http://dx.doi.org/10.1023/A:1021271517844> doi: 10.1023/A:1021271517844
- Hancock, M., ten Cate, T., & Carpendale, S. (2009). Sticky tools: Full 6dof force-based interaction for multi-touch tables. In *Proceedings of the acm international conference on interactive tabletops and surfaces* (pp. 133–140). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/1731903.1731930> doi: 10.1145/1731903.1731930
- Henderson, J., Pizarro, O., Johnson-Roberson, M., & Mahon, I. (2013). Mapping submerged archaeological sites using stereo-vision photogrammetry. *International Journal of Nautical Archaeology, 42(2)*, 243–256.
- Herrlich, M., Walther-Franks, B., & Malaka, R. (2011). Integrated rotation and translation for 3d manipulation on multi-touch interactive surfaces. In *Proceedings of the 11th international conference on smart graphics* (pp. 146–154). Berlin, Heidelberg: Springer-Verlag. Retrieved from <http://dl.acm.org/citation.cfm?id=2032567.2032589>
- Kazhdan, M., & Hoppe, H. (2013). Screened poisson surface reconstruction. *ACM Transactions on Graphics (TOG), 32(3)*, 29.
- Kim, Y.-S., Kesavadas, T., & Paley, S. M. (2006). The virtual site museum: A multi-purpose, authoritative, and functional virtual heritage resource. *Presence: Teleoperators and Virtual Environments, 15(3)*, 245-261. Retrieved from <https://doi.org/10.1162/pres.15.3.245> doi: 10.1162/pres.15.3.245
- Kulik, A. (2016). *User interfaces for cooperation* (PhD thesis). Bauhaus-Universität Weimar.
- Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., & Fröhlich, B. (2011, December). C1x6: a stereoscopic six-user display for co-located collaboration in shared virtual environments. *ACM Trans. Graph., 30(6)*, 188:1–188:12. Retrieved from <http://doi.acm.org/10.1145/2070781.2024222> doi: 10.1145/2070781.2024222
- Kunert, A., Kulik, A., Beck, S., & Fröhlich, B. (2014). Photoportals: Shared references in space and time. In *Proceedings of the 17th acm conference on computer supported cooperative work & social computing* (pp. 1388–1399). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/2531602.2531727> doi: 10.1145/2531602.2531727

- Lawonn, K., Trostmann, E., Preim, B., & Hildebrandt, K. (2017, Jan). Visualization and extraction of carvings for heritage conservation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 801-810. doi: 10.1109/TVCG.2016.2598603
- Liu, J., Au, O. K.-C., Fu, H., & Tai, C.-L. (2012, September). Two-finger gestures for 6dof manipulation of 3d objects. *Comp. Graph. Forum*, 31(7pt1), 2047-2055. Retrieved from <http://dx.doi.org/10.1111/j.1467-8659.2012.03197.x> doi: 10.1111/j.1467-8659.2012.03197.x
- Maimone, A., & Fuchs, H. (2011). Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras. In *Proc. of ismar 2011* (pp. 137-146). Washington, DC, USA: IEEE Computer Society. Retrieved from <http://dx.doi.org/10.1109/ISMAR.2011.6092379> doi: 10.1109/ISMAR.2011.6092379
- Marretta, A. (2007). Forma, funzione e territorio nell'arte rupestre camuna: il caso delle figure oritomorfe. In *Valcamonica symposium* (Vol. 2007, p. 1).
- Marretta, A. (2013). Age of the heroes: a brief overview of valcamonica rock-art during the iron age (i millennium bc) valcamonica world heritage. *Adoranten*, 75.
- Martinet, A., Casiez, G., & Grisoni, L. (2010). The effect of dof separation in 3d manipulation tasks with multi-touch displays. In *Proceedings of the 17th acm symposium on virtual reality software and technology* (pp. 111-118). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/1889863.1889888> doi: 10.1145/1889863.1889888
- McCrae, J., Mordatch, I., Glueck, M., & Khan, A. (2009). Multiscale 3d navigation. In *Proceedings of the 2009 symposium on interactive 3d graphics and games* (pp. 7-14). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/1507149.1507151> doi: 10.1145/1507149.1507151
- Olson, B. R., Placchetti, R. A., Quartermaine, J., & Killebrew, A. E. (2013). The tel akko total archaeology project (akko, israel): Assessing the suitability of multi-scale 3d field recording in archaeology. *Journal of Field Archaeology*, 38(3), 244-262.
- Pauly, M., Gross, M., & Kobbelt, L. P. (2002). Efficient simplification of point-sampled surfaces. In *Proceedings of the conference on visualization'02* (pp. 163-170).
- Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., & Chamzas, C. (2007). Methods for 3d digitization of cultural heritage. *Journal of cultural heritage*, 8(1), 93-98.
- Pfister, H., Zwicker, M., van Baar, J., & Gross, M. (2000). Surfels: Surface elements as rendering primitives. In *Proceedings of the 27th annual conference on computer graphics and interactive techniques* (pp. 335-342). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. Retrieved from <http://dx.doi.org/10.1145/344779.344936> doi: 10.1145/344779.344936
- Pierdicca, R., Frontoni, E., Malinverni, E. S., Colosi, F., & Orazi, R. (2016). Virtual reconstruction of archaeological heritage using a combination of photogrammetric techniques: Huaca arco iris, chan chan, peru. *Digital Applications in Archaeology and Cultural Heritage*, 3(3), 80-90.
- Priuli, A. (1985). Incisioni rupestri della val camonica.
- Reunanen, M., Díaz, L., & Horttana, T. (2015, February). A holistic user-centered approach to immersive digital cultural heritage installations: Case vrouw maria. *J. Comput. Cult. Herit.*, 7(4), 24:1-24:16. Retrieved from <http://doi.acm.org/10.1145/2637485> doi: 10.1145/2637485
- Rogers, Y., & Lindley, S. (2004). Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers*, 16(6), 1133-1152.
- Schollmeyer, A., Schneegans, S., Beck, S., Steed, A., & Froehlich, B. (2017). Efficient hybrid image warping for high frame-rate stereoscopic. *IEEE Transactions on Visualization and Computer Graphics*.
- Scott, S. D., Sheelagh, M., Carpendale, T., & Inkpen, K. M. (2004). Territoriality in collaborative tabletop workspaces. In *Proc. of cscw 2004* (pp. 294-303). ACM Press. Retrieved from <http://doi.acm.org/10.1145/1031607.1031655> doi: 10.1145/1031607.1031655
- Smith, C. (2014). *Encyclopedia of global archaeology*. Springer Reference.
- Soler, F., Melero, F. J., & Luzón, M. V. (2017). A complete 3d information system for cultural heritage documentation. *Journal of Cultural Heritage*, 23(Supplement C), 49 - 57. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1296207416302576> doi: <https://doi.org/10.1016/j.culher.2016.09.008>
- Strothoff, S., Valkov, D., & Hinrichs, K. (2011). Triangle cursor: Interactions with objects above the tabletop. In *Proceedings of the acm international conference on interactive tabletops and surfaces* (pp. 111-119). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/2076354.2076377> doi: 10.1145/2076354.2076377
- Thwaites, H. (2013). Digital heritage: What happens when we digitize everything? In E. Ch'ng, V. Gaffney, & H. Chapman (Eds.), *Visual heritage in the digital age* (pp. 327-348). London: Springer London. Retrieved from https://doi.org/10.1007/978-1-4471-5535-5_17 doi: 10.1007/978-1-4471-5535-5_17
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004, April). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600-612. Retrieved from <http://dx.doi.org/10.1109/TIP.2003.819861> doi: 10.1109/TIP.2003.819861
- Webel, S., Olbrich, M., Franke, T., & Keil, J. (2013, Oct). Immersive experience of current and ancient reconstructed cultural attractions. In *2013 digital heritage international congress (digitalheritage)* (Vol. 1, p. 395-398). doi: 10.1109/DigitalHeritage.2013.6743766
- White, R. H. (2013). Resolving the carving: The application of laser scanning in reconstructing a viking cross from neston, cheshire. In E. Ch'ng, V. Gaffney, & H. Chapman (Eds.), *Visual heritage in the digital age* (pp. 33-41). London: Springer London. Retrieved from https://doi.org/10.1007/978-1-4471-5535-5_3 doi: 10.1007/978-1-4471-5535-5_3
- Zeppelzauer, M., Poier, G., Seidl, M., Reinbacher, C., Breiteneder, C., Bischof, H., & Schuster, S. (2015). Interactive segmentation of rock-art in high-resolution 3d reconstructions. In *Digital heritage, 2015* (Vol. 2, pp. 37-44).