A View from the Hill: Where Cross Reality Meets Virtual Worlds

Abstract

This paper reports experience in developing a cross reality system which enables a user to be present and aware of both the virtual environment and the real world at the same time. In so doing the challenge of the vacancy problem is addressed by lightening the cognitive load needed to switch between realities and to navigate the virtual environment. Evaluation of the usability, system performance and value of the system are undertaken in the context of a cultural heritage application. Users are able to compare a reconstruction of an important 15th Century chapel with its present day instantiation.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Interfaces—Mixed/Augmented Reality

Keywords: Cross reality, digital heritage, immersive interaction

1 Introduction

This research centers around the design, development & evaluation of a hardware & software platform which allows its user to observe & move around their Real World (RW) environment whilst wearing a wide field of view (FOV), stereoscopic 3D, Head Mounted Display (HMD) which allows them to alternatively view an immersive Virtual Reality (VR) environment from the equivalent vantage point. This is achieved by combining a head-tracked HMD, webcams, an indoor positioning system (IPS) & a 3D game engine, into a mobile *cross reality* (XR) interface.

One of the distinguishing features of XR is that, by linking real & virtual environments more closely, it mitigates the 'vacancy problem': *"the noticeable & profound absence of a person from one world, either real or virtual, while they are participating in the other"*, which arises *"because people do not currently have the means to be in more than one place (reality) at a time"* [Lifton 2007].

Previous XR research approached the vacancy problem by integrating sensor/actuator networks into the environments, such that actions in one could manifest in the other, however direct visual engagement with the virtual environment was only possible from static interfaces at pre-determined locations within the real environment [Lifton 2007] [Dublon et al. 2011]. The platform discussed in this document addresses this shortcoming by providing a mobile interface for visual engagement with both environments of a XR system, allowing the user to transition between viewing their real environment & a virtual environment at any time while maintaining the freedom to move around them, multiplexing visual stimuli from their real surroundings & from a parallel, virtual 'mirror world' [Gelernter 1993].

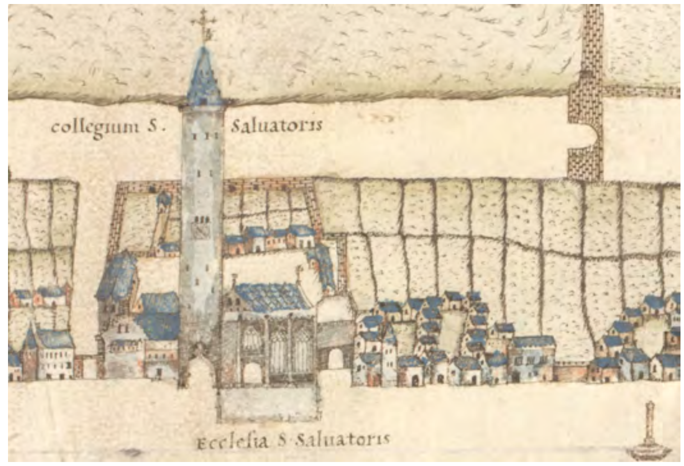

Figure 1: *Geddy View of St Salvator's Chapel 1552.*

2 Context and Related Work

The position of XR in relation to other alternate realities studied by Computer Science can be visualised using Milgram & Kishino's *virtuality continuum* (figure 2) [Milgram and Kishino 1994] that stretches from an entirely real environment at one extreme to an ontologically parallel but entirely virtual environment [Qvortrup 2002] at the other. The explanation herein distinguishes between environments themselves (depicted in figure 4 by solid ellipses) & where the stimuli that the user is perceiving originate from (depicted by dashed ellipses).

Ontology - The core meaning within computer science is a model for describing the world that consists of a set of types, properties, and relationship types. There is also generally an expectation that the features of the model in an ontology should closely resemble the real world (related to the object).

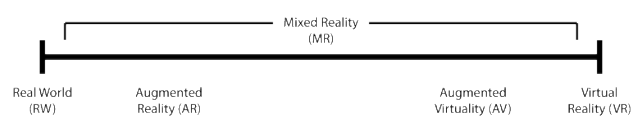

Figure 2: *Milgram & Kishino's virtuality continuum.*

Of particular importance is to appreciate the distinction between a XR system & an *augmented reality* (AR) system, as both concepts involve user engagement with both real & virtual content. Whilst an AR system features a single environment, comprised of the user's RW overlain by some virtual content, with the user perceiving stimuli from this single augmented environment, a XR system instead features two discrete environments, one real & the other virtual, each complete unto itself.

Whilst AR falls within the realms of Mixed Reality (MR), a XR system can be considered as occupying the two extremes of the continuum outwith the MR region. However, XR systems that allow simultaneous interaction with both of their constituent environments blur this definition; using a XR platform such as that discussed in this document, a user can transition between perceiving stimuli from each of these environments in a manner that allows them to engage with each environment without becoming wholly vacant from the other.

Figure 3: *The two environments that comprise a XR system.*

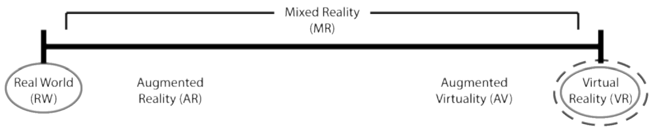

Figure 4: *A XR system with the user attending to VR stimuli.*

The virtuality continuum is here considered to be analogous to the *locus of attention* axis of Waterworth & Waterworth's *three dimensions of virtual experience* model [Waterworth and Waterworth 2001]; the combination of these models is shown by figure 5. In this model, locus of attention represents the environment where the stimuli that the user is perceiving originate from; focus of attention represents the balance between conceptual/abstract reasoning & perceptual/concrete processing, where complex conceptual reasoning results in little attention being paid to processing environmental percepts (whether originating from real or virtual stimuli) thus reducing presence[1] in that environment toward its antithesis – absence[2]; and sensus of attention represents the level of conscious arousal (or 'wakefulness' [Laureys et al. 2009]) of the user, whether directed toward percepts originating from real stimuli, virtual stimuli, a mix, or not directed toward any percepts in the case of completely 'absent' conceptual reasoning.

2.1 The Case for Mobile XR

A XR system that presents the user with visual stimuli from both its constituent environments (RW & VR) allows that user to engage with both real & virtual content in a manner that is similar to, but has a number of advantages over, a traditional AR system;

- the XR system is less critical of registration (the accurate positioning/alignment) between real & virtual, as the virtual objects are seen as part of a larger virtual environment instead of being rendered atop a view of the real environment;
- the XR system can make use of existing VR content without the overhead of decanting/extracting a subset of the virtual components into an AR framework (e.g. manually selecting which objects within the VR environment are to be displayed over the RW environment);
- the use of a complete VR environment allows the virtual content to be more encompassing & immersive, as presenting a complete VR environment allows total control over lighting, shadows, reflections, particle effects, etc. which would be difficult or impossible for an AR platform to render atop a view of a RW environment.

[1]**Presence** in this context is defined as a state of heightened perceptual processing of environmental stimuli (*"a psychological focus on direct perceptual processing"* [Waterworth and Waterworth 2001]) accompanied by lessened conceptual reasoning, whether these environmental stimuli originate from a real environment, a virtual environment, a mixed reality environment, or even from multiple discrete environments.

[2]**Absence** is defined as *"a psychological focus on ... conceptual processing"* [Waterworth and Waterworth 2001].

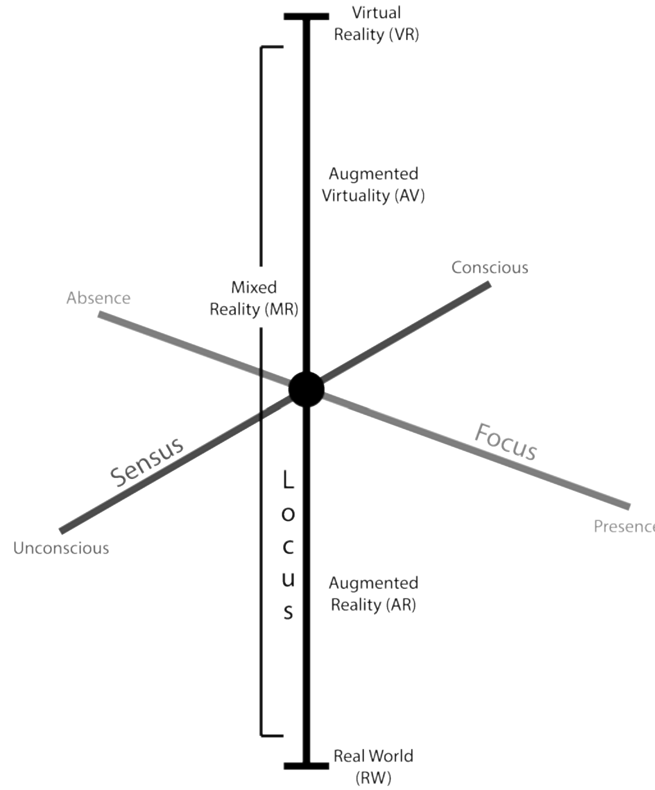

Figure 5: *The combined virtuality continuum/three dimensions of virtual experience model.*

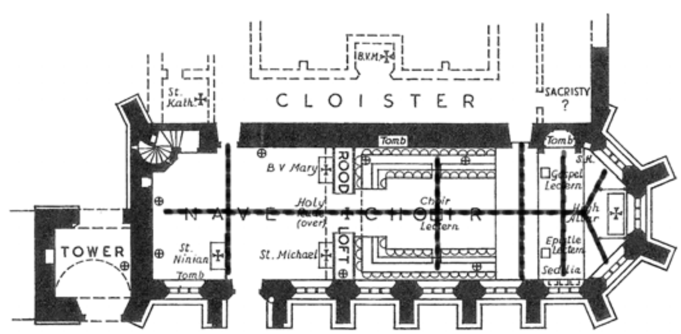

Figure 6: *Floor plan of St Salvator's chapel, with IPS routes.*

Thus, such a XR platform is well suited to situations in which interaction with both real & virtual visual stimuli is required & where one or more of the following hold true;

- in lieu of accurate registration between real & virtual, there is a strong focus on the virtual environment's atmosphere & immersion [De Amicis et al. 2009];
- the visual differences between real & virtual environments are so substantial that an AR system would resort to augment (&/or diminish [Mann and Fung 2002]) almost the whole RW view. While AR *"smears an informational coating over real space"* [Andersen], XR presents a complete, discrete virtual environment. AR is beneficial where one wishes the juxtaposition of virtual objects upon what is already present in the RW environment, however VR is better suited to situations where one wishes to present a complete virtual alternative.

Figure 7: *Recent View of St Salvator's from Quad.*

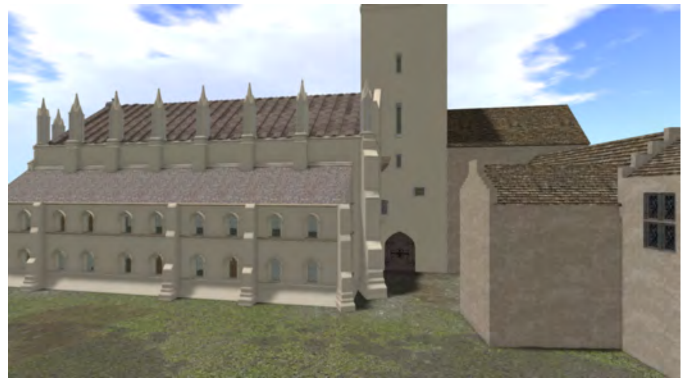

Figure 8: *Reconstruction of St Salvator's Circa 1450.*

3 Example Application - Cultural Heritage

The field of cultural heritage has seen widespread applications of both AR [Kim et al. 2009][Wojciechowski et al. 2004] [Okura et al. 2006] [Magnenat-Thalmann et al. 2008] [Vlahakis et al. 2001] [Okura et al. 2011] [Papagiannakis et al. 2004] [Magnenat-Thalmann et al. 7] [Papagiannakis et al. 2005] [Papagiannakis and Magnenat-Thalmann 2007] [Papagiannakis and Magnenat-Thalmann 2009] [Taketomi et al. 2011] [Walczak and White 2003] & VR [De Amicis et al. 2009] [Roussou et al. 2003] [Cabral et al. 2007] [Roussou 2002] [Christou et al. 2006] [Willmott et al. 2001] [Tzortzaki 2002]. AR has been used to add artefacts, actors & reconstructed architecture to views of present day sites that bear traces of their original status, whilst VR has been used to host more complete reconstructions of entire buildings & settlements for interaction via screen, HMD & CAVE, including where the present day site bears no evidence of the past status or is inaccessible for some reason (due to latter development, change in landscape, etc.).

In situations where VR content exists in cultural heritage contexts, it is experienced from a static position that causes both spatial & temporal separation from the RW location that it relates to; in order to perform comparisons between RW & VR content, users must interact with one & *subsequently* the other. A mobile XR platform will allow VR content in cultural heritage contexts to be experienced in tandem with the real site (where accessible), combining the immediate juxtaposition of real & virtual content of AR with the immersive & atmospheric qualities of HMD based VR, all without requiring alterations to the VR content (for example to make it compatible with an AR framework).

3.1 St Salvator's Chapel Reconstruction

Founded in 1450 but internally stripped of its medieval fittings during the Protestant Reformation (1517 - 1648), St Salvator's chapel in St Andrews looks markedly different in the present day than it did upon its completion. An existing VR reconstruction of the chapel as it stood in the period 1450-1460 & the marked differences between the internal appearance of the VR building & the current building (including the replacement of the original stone roof with a wooden one & drastically different dividing of the internal space) make this chapel an ideal candidate within the context of cultural heritage for a mobile XR system to be applied. Figure 6 shows the 1450-1460 layout of the chapel (including the paths that the IPS has been prepared upon).

This project has created a reconstruction of St Salvators Chapel as it was built and furnished for Bishop James Kennedy between 1450 and 1460. The chapel was of the greatest significance for the new architectural ideas that it introduced into Scotland, at a time when Scotland was particularly open to external artistic influences. However, although the shell of the chapel survives and remains in use, it has lost its vault, its window tracery and its liturgical furnishings, and it now requires specialist skills to appreciate the quality of its original state.

Collaboration between architectural and art history and computer science scholarship has created a 3D reconstructions of the chapel located within a virtual world. The virtual chapel enhanced with lighting, sound and movement may be explored from a variety of physical and temporal locations through an avatar. The architectural, liturgical and spatial analysis allows our understanding of the history of the Chapel as a living building to be enormously enhanced by experiencing the building in its original context.

St Salvator's college was founded on 27 August 1450 by Bishop James Kennedy, who played a leading role in Scottish and international politics. St Salvators was to be open to students who were prepared to live within the college. In this St Salvators was the first college in Scotland to place the education of students so firmly at the heart of its role. It was dedicated for worship in 1460. The chapel is an aisle-less rectangle with a three-sided east apse. Deeply projecting three-stage buttresses define the bays, which are now capped by pinnacles of 18612.

The windows which occupy the full space available between the buttresses, no longer reflect their original forms. The main entrance to the chapel was through a doorway in the second bay from the west of the south flank, which is covered by a vaulted porch between the buttresses. Two doorways on the north side presumably opened into a lost sacristy and treasury range.

The interior of the chapel is known to have been covered by a stone vault, which is assumed to have been of pointed barrel form with a decorative pattern of ribs, like the small vault over the south porch. The interior is now covered by an inappropriate timber roof.

St Salvators chapel is considered the first Scottish example of a church planned with an aisle-less rectangular main body terminating in a polygonal eastern apse, a type that was to have a long future for a range of Scottish church types. Such chapels were common in university colleges in France, and since Bishop Kennedy had a highly placed kinsman in the university of Paris and drew many ideas for the organisation of his college from that universitys constitution, it is reasonable to assume that he also drew some of his ideas for the architecture of his chapel from there. On this basis, St Salvators must be seen as an outstandingly important channel for the introduction into Scotland of new architectural ideas from France. The new architecture made a significant statement in its

Figure 9: *Non Player Characters Celebrate High Mass.*

Figure 10: *The virtual and the real in the same time and place.*

Scottish context.

Collaboration between researchers in Art History, Classics and Computer Science has enabled a historically accurate reconstruction of the Chapel to be created. On the combined evidence of a highly detailed late medieval inventory and of the architecture itself, it has been possible to show how the chapel was furnished internally with altars, choir stalls, lecterns, screens, stained glass and wall paintings.

The reconstruction of the chapel involved both the mental reconstruction of modified and lost features, and the establishment of the range of ways in which buildings that represent a spirituality alien to modern students were intended to function. As such it offers an invaluable academic discipline for those involved in the reconstruction, providing eminently practical ways of testing theories and assumptions. It is then of the greatest value for conveying more widely the understanding that has been gained. The development of a cross reality system which enables comparison between the real and virtual chapel in the same time and place aims to further enhance the value of the reconstruction.

4 Mirrorshades Design and Implementation

A high level architectural overview of our mobile XR platform is shown in Fig. 11, dubbed Mirrorshades[3]. The Head mounted display enables the user to experience both the real world and the virtual world in stereoscopic 3D. It receives both the virtual and real world feeds from the mobile client. The head mounted display track user head movements and relays these to the client so that the virtual view can be appropriately adjusted. The indoor positioning system tracks the users location and forwards this to the mobile client. The mobile client runs the virtual world and controls the position of the user's proxy within it. The Video cameras are mounted on the HMD and provide their feed to the client. The client recieves input from the controller and uses this to choreograph the the sending of real and virtual video feeds to the head mounted display.

- an Oculus Rift DK1 HMD, including a 9-axis (3dof rotational) head tracker sampling at 1000Hz & mounted with a stereo camera solution comprising 2x Logitech C310 webcams modified with M12 lens mounts & 2.1mm lenses to provide approximately 87 degrees horizontal FOV of the RW environment (see figure 13);
- a USB battery pack, to power the HMD;
- a small laptop computer, with an Intel i7-3632QM processor, Nvidia GT 650M graphics card & 16GiB system memory;
- an Android smartphone, running Android 4.4.4;
- an Xbox 360 wireless controller, with USB receiver.

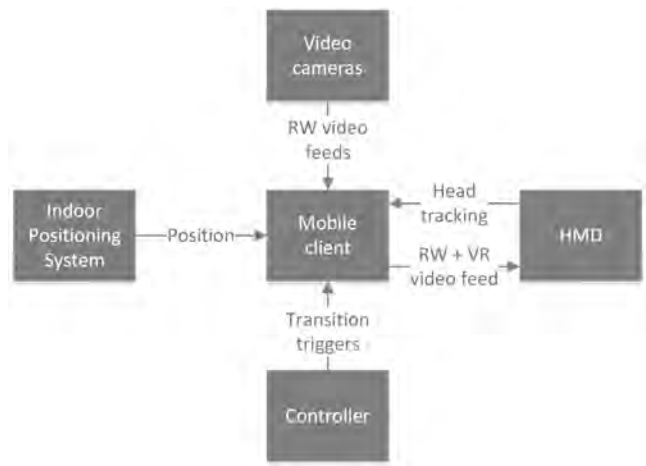

Figure 11: *Overview of the Mirrorshades platform.*

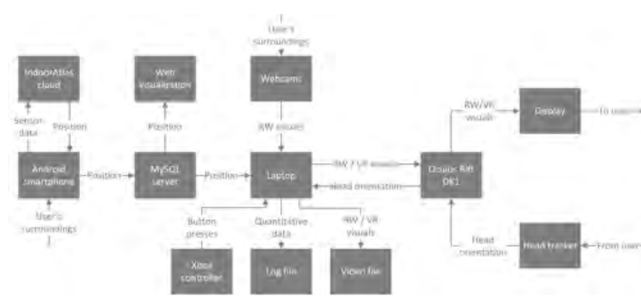

Figure 12: *Implementation of Mirrorshades platform.*

The Logitech C310 & has a resolution of (1280960), the refresh rate is lower (30hz). They are held in place by 3D printed clips (red) with the cameras epoxyd to thermoplastic (white) so they can be adjusted via the nuts & bolts with rubber washers.

An overview of the implementation of the Mirrorshades platform design for use in the chapel investigations is shown in Fig. 12. The Unity application hosts the VR representation of the chapel & takes in feeds from both webcams, the HMD head tracker & the

[3]**Mirrorshades: The Cyberpunk Anthology** (1986) is a defining cyberpunk short story collection, edited by Bruce Sterling.

Figure 13: *HMD with stereo camera solution.*

Xbox controller. It also polls the database server for the most recent position data. All of these inputs are combined together to form the visual output for the HMD to display to the user.

As the user moves their head, the visuals that are presented to them upon the HMD's display change accordingly; the RW visuals change due to the webcams being physically fixed to the HMD & the VR visuals change due to data from the head tracker being used to change the orientation of the in game 'cameras' accordingly.

As the user changes their position by walking, the visuals that are presented to them upon the HMD's display also change accordingly; again the RW visuals change due to the webcams' position upon the HMD whilst the VR visuals change due to the user's position, as reported by the smartphone & the IndoorAtlas solution, being used to move the position of the in game cameras to the equivalent position within the VR representation.

As the user presses buttons or pulls triggers upon the Xbox controller, the visuals that are presented to them upon the HMD's display transition between RW & VR in different styles depending upon which button/trigger was activated.

Figure 14: *Chapel pews, greater in number in present day.*

5 Between the Real to the Virtual

The novel aspect of the Mirrorshades platform is the ability it imparts upon its user to switch their locus of attention between equivalent vantage points in RW & VR environments whilst walking around. This is achieved by the user performing transitions between RW visual stimuli & VR visual stimuli, both presented via their HMD. This extends existing XR research by allowing the user to engage with the visual stimuli of the VR component of a XR system from any position & at any time.

Figure 15: *Chapel ceiling, different construction in present day.*

In order to achieve the highest quality of experience with this style of interaction with XR systems, it is vital to determine how best to implement these transitions; that is, to mitigate the increased cognitive load (manifesting as increased conceptual reasoning & reduced perceptual processing, see section **??**) required to comprehend these transitions, as this increased cognitive load will detract from engagement with the environments & reduce the user's willingness to perform these transitions.

Whilst some researchers support the notion that in systems where more than one environment competes for the user's locus of attention there is an 'all or nothing' Gestalt switch between awareness of one environment & the other [Slater 2002], which would result in a substantial increase in cognitive load upon each transition, Mirrorshades has been developed in support of the contrary opinion; that switching locus of attention from the stimuli of one environment to those of another does not completely overrule the user's awareness of the former, that both environments can be perceived at the same time (albeit one to a lesser extent) [Ijsselsteijn 2001] & that when engaging with VR content a user's focus can even be said to typically be *shared* between VR & RW [Waterworth and Waterworth 2001].

This latter position is particularly apt for situations wherein the RW & VR environments share the same fundamental layout & dimensions, as those in a XR system often do, as the inherent familiarity between the two environments reduces the cognitive load associated with transitioning between them. Furthermore, the notion of experience of presence as changing continually from moment-to-moment [Heeter 2003] [Ijsselsteijn et al. 1998] lends confidence to the successful mitigation of the cognitive load associated with these transitions to manageable levels. One might even liken this 'switching' between RW & VR to the 'cycling through' behaviour observed in users of virtual communities, which stemmed from the 'window' concept of modern computer operating systems [Turkle 2004].

However, no matter how smooth the transition the process is expected to always result in some heightened cognitive load, a temporary *break in presence*[4] (BIP), as the user comes to terms with

[4]The definition of **break in presence** adopted herein is the second from Waterworth & Waterworth [Waterworth and Waterworth 2001] (p205): a movement along the focus axis away from presence in the real or a virtual environment & toward absence. This differs to Slater & Steed's original definition in [Slater and Steed 2000] as they considered presence only in terms of attending to stimuli from a virtual environment, with a break in presence

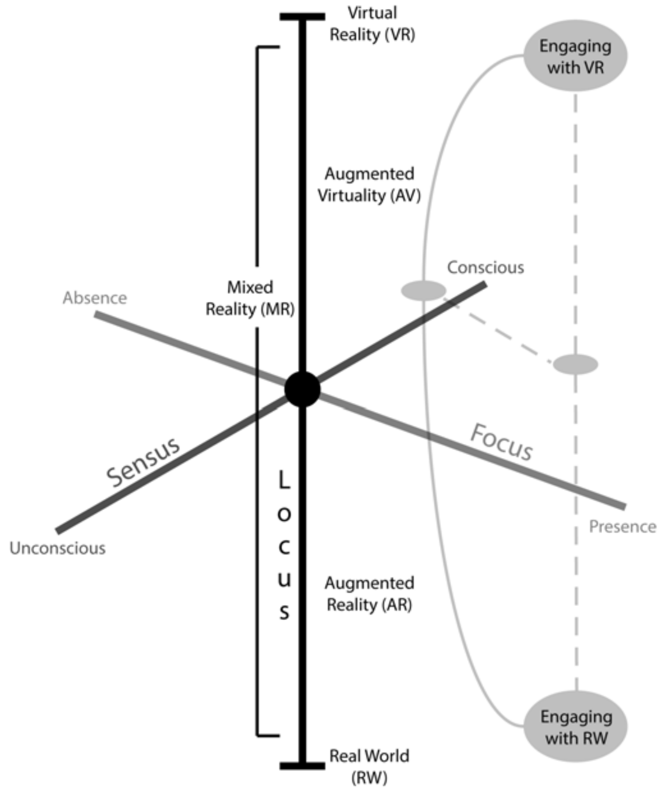

Figure 16: *Operation of the Mirrorshades platform represented upon the combined model.*

Figure 17: *Hard Switch*

Figure 18: *Switch with linear interpolation*

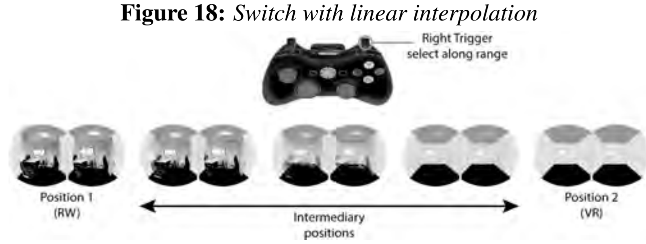

Figure 19: *Analogue selectable opacity*

the new environment presented to them & comprehends its relation to the other environment that they were just perceiving.

Several different transition methods have been developed & this investigation identifies & quantifies preferences toward them, to infer which approaches to transitioning between RW & VR visual stimuli are more or less appropriate for the different situations that arise where a platform like Mirrorshades may be deployed. In particular, it is hypothesized that there is a strong correlation between participant movement (or lack thereof) & choice of particular transition style.

Visualised using the combined model (see section **??**) as figure 16, these transitions are an oscillation along the locus axis, between a RW environment at one position & a VR environment at the other.

Heightened cognitive load required to comprehend a transition is a temporary movement upon the focus axis from presence toward absence (a BIP). With the ability of a wide FOV, stereoscopic 3D, head-tracked HMD to produce immersive VR visual stimuli that require fairly limited cognitive processing & our inherent ability to engage with our RW surroundings without significant cognitive load, focus is expected to be high (toward the presence extremum) when attending to stimuli from either RW or VR.

Sensus is expected to be largely task dependent, however when performing a task that involves actively engaging with the visual stimuli from either/both of RW or VR it is expected to be high (toward the conscious extremum). Upon triggering a transition, sensus is expected to increase, as the user centres their attention upon relating the visual stimuli from the new environment to those they were just perceiving from the other environment.

Attending to visual stimuli from the RW environment via the webcams is required for the user to safely move around. Delay in the IPS reporting their position & inaccuracies in these position data (see figure **??** for a set of example position data) mean that moving around while attending only to visual stimuli from the VR environment would not be safe for the user, even with unchanging RW obstacles with perfectly accurate representations in the VR environment. Furthermore it is actually likely that RW obstacles will not have equivalent VR representations, such as in a scenario where XR is used to compare & contrast changes to a building's interior over extended periods of time (such as with the chapel investigations).

Thus the HMD displays the feeds from the webcams as default & the user must trigger transitions to view the VR environment by pressing a button or pulling a trigger on the controller. Releasing the button/trigger causes the webcam feeds to be displayed again.

Four modes of switching were considered, hard switch, switch with linear interpolation, selectable opacity and automated switching. Automated switching was discarded as it did not meet user needs.

---

as a Gestalt switch to instead attending to stimuli from the real environment. Waterworth & Waterworth's model considers presence in terms of attending to stimuli from either the real *or a virtual* environment, with a break in presence representing absence in the sense of heightened conceptual load & the resultant reduced perceptual processing of environmental stimuli originating from *either* the real or a virtual environment. This definition better fits the situation invoked by the Mirrorshades platform, which is concerned with intentionally & willingly switching engagement between stimuli from both real & virtual environments, rather than engaging with stimuli from only a virtual environment in a scenario where stimuli from the real environment are considered a 'distraction'.

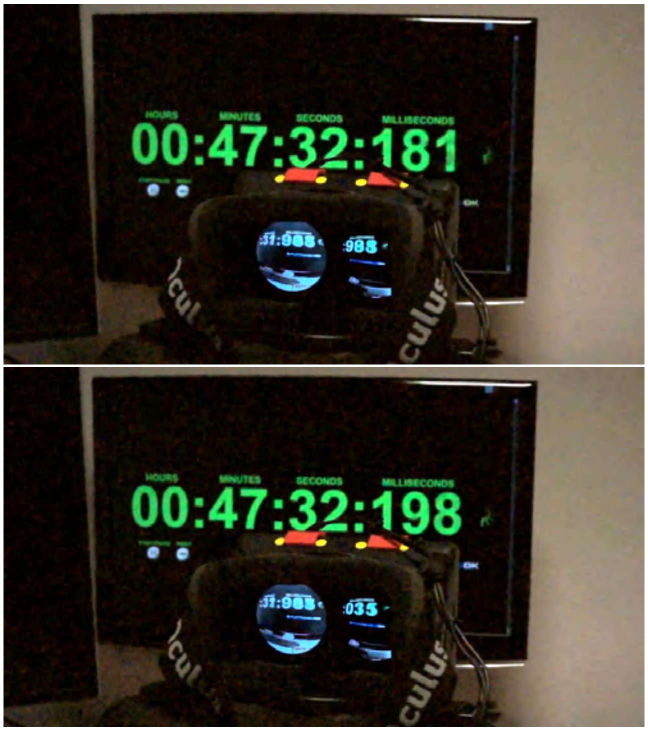

Figure 20: *Measurement of latency introduced by cameras*

Hard switch The user presses & holds the [A] button on the controller to switch the visual stimuli displayed by the HMD from RW to VR. When the [A] button is released, the visual stimuli displayed by the HMD switch back from VR to RW. This is a 'hard' or 'immediate' switch with no fading or transition effect. Figure 17 illustrates this scenario.

Switch with linear interpolation The user presses & holds the [B] button on the controller to switch the visual stimuli displayed by the HMD from RW to VR. When the [B] button is released, the visual stimuli displayed by the HMD switch back from VR to RW. This switch fades between RW & VR visual stimuli using linear interpolation on the opacity of the game objects that the webcam feeds are rendered upon. Figure 18 illustrates this scenario.

Analogue selectable opacity The user pulls the right analogue trigger ([RT]) on the controller, where the position of the trigger maps directly to the opacity of the game objects that the webcam feeds are rendered upon. The user can choose to stop at any intermediary position that suits their needs, keeping the level of opacity of the webcam feeds at that position, as well as controlling the rate at which the visual stimuli from the RW environment fade (by changing how quickly they change their depression of the trigger). Pulling the trigger all the way in displays only visual stimuli from the VR environment, while releasing it completely displays only visual stimuli from the RW environment. The number of intermediary positions is limited only by the resolution of the trigger & the encoding of the value.

This method allows the user to superimpose VR visual stimuli upon RW visual stimuli. This is similar, but not identical, to AR, as instead of displaying a small number of virtual objects upon the user's view of their RW environment, a complete VR environment is superimposed upon the user's view of their RW environment. Figure 19 illustrates this scenario.

6 System and User Evaluation

The contributions of this paper flow from taking cross reality out of the lab and applying it to a concrete cultural heritage scenario. We describe a system which has been engineered to address the vacancy problem and enable comparison between the virtual and the real. The goals of the evaluation are threefold. Firstly, to asses the system performance of Mirrorshades. How accurately is the user tracked, how responsive is the system and the frame rate is sustained. Secondly, to compare methods used to switch between realities and thirdly to determine the value that users place on the system.

Evaluating users' preferences toward different methods of transitioning between visual stimuli in different situations pertains to studying their reactions & responses to ascertain the effect upon their focus of attention, concepts which are largely psychological in nature & highly subjective [Ijsselsteijn 2001]. Thus, subjective measures are important in this evaluation particularly feedback from users. However, objective data has also been collected & cross referenced with the subjective data in attempts to support or contradict any relationships that are identified.

It is hypothesized that a manner of transitioning between visual stimuli which results in a less severe BIP will be preferable to a manner of transitioning which results in a worse BIP. As focus in the Waterworth model is most closely related to presence in the VR literature [Waterworth and Waterworth 2001], one of the subjective measures that will be used in this evaluation is an established presence measure, to try to capture the behaviour of the user's position upon the focus axis.

For these experiments, the HMD was worn upon the head of the participant & connected to the laptop, battery pack & wireless receiver worn in a satchel. The smartphone is held in the left hand & the Xbox controller is held in the right hand (all of the buttons & triggers used for these experiments are on the right hand side of the controller, designed to be activated with only the right hand). Participants were encouraged to take their time explore St Salvator's, walking from the nave to the East end, to examine the Kennedy tomb and to return to the West end.

6.1 Data Capture

The Unity app logs the following quantitative data each frame to a tab separated variable (.tsv) file;

- <frame number>
- <timestamp> - according to the laptop's internal clock
- <original_position> - the position as a Unity Vector3 where the user begins the experiment
- <position> - the position as a Unity Vector3 where the user is on this frame
- <delta_x> & <delta_z> - the difference in the x & z axes between <original_position> & <position> on this frame
- <left_rotation> & <right_rotation> - the orientations as Unity Quaternion of the two Unity camera game objects
- <base_oapcity> - the maximum opacity of the game objects upon which the webcam feeds are rendered (see section **??**)

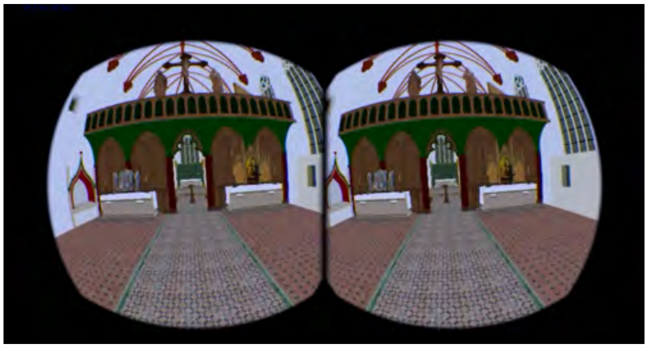

Figure 21: *View of Virtual St Salvator's*

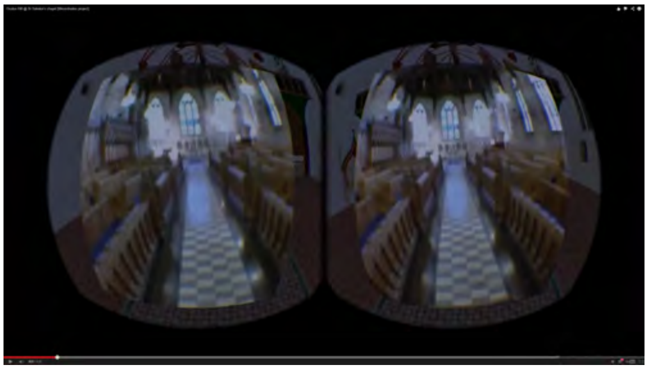

Figure 22: *View of Real through Cameras attached to Rift*

- `<left_opacity>` & `<right_opacity>` - the opacity on this frame of the game objects upon which the webcam feeds are rendered
- `<auto_tick>` - whether a periodic switch is in progress (see section **??**)
- `<auto_duration>` & `<auto_spacing>` - the interval & duration values of the periodic hard switching
- `<framerate>` - an estimate of the current frame rate (frames per second)
- `<A_button>`, `<B_button>` & `<right_trigger>` - the current values of these inputs on the controller

An example line of this output;

```
420 08-05-2014 12-34-36-257 (3.4, 1.0, -8.3)
    (0.3, 1.0, -8.3) 3.153522 0.0001955032
  (-0.1, -0.7, -0.1, 0.7) (-0.1, -0.7, -0.1,
 0.7) 1 1 1 False 0 0 39.57977 False False 0
```

These data reveal relationships between various different metrics & the choice of transition methods. For example, it is expected that participants will perform short transitions to VR or transitions to a mix of RW & VR when moving & perform longer transitions to VR when stationary. This kind of relationship will support or contradict the subjective data collected through questionnaire & interview.

During experiments, the video feed being displayed by the HMD will be recorded & the user will be recorded using a video camera (both video & audio). The video of the HMD graphics will be used in comparison with the quantitative data, while the video & audio recording of the user will provide objective insight into their behaviour. Video from these feeds can be viewed here[5]

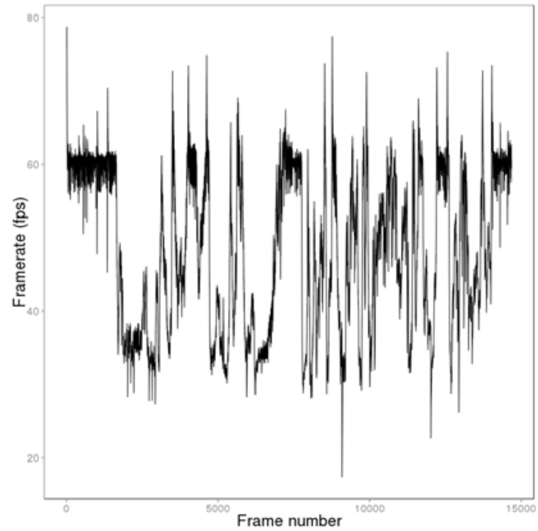

Figure 23: *Oculus Rift frame rate during use in St Salvator's*

6.2 System Evaluation

The latency introduced by the C310 webcams was measured. The Rift, with the lenses removed was placed, facing a LCD monitor displaying a timer from flatpanels. A camera was placed behind such that it could see both the monitor & the Rifts screen, then the sensitivity on the camera was adjusted so that it could record 50fps video with a 1/4000th shutter speed.

The monitor & the Rift were both refreshing at 60fps, each frame lasting 16.67ms, whilst a 1/4000 shutter speed on the camera meant that the shutter was open for 0.25ms. The response time of the monitor (quoted by the manufacturer as 8ms GTG) was evidently much higher than that of the Rift, as the tenths & even hundredths digit on the monitor was usually legible in each frame of the video whereas on the Rift the hundredths & thousandths digits were always illegible. So the video was inspected frame-by-frame looking for adjacent frames where a transition from one tenth digit to the next was good enough to read on the Rift & the hundredths/thousandths digits were good enough to read on the monitor, such as the pair in Fig 20

From these it can be inferred that the tenths digit on the Rift screen (right eye) changed from 9 to 0 sometime between 181 & 198 on the monitor, meaning a latency of between 181ms & 198ms. Out of 11 pairs of frames like this, 7 pairs showed this 181-198ms latency, whilst 4 pairs showed 198-215ms. The latency was confirmed by taking some still photos with the same 1/4000th shutter speed, they all showed the same 181-215ms latency, however as timing shots to get legible digits was difficult it was more reliable to video at 50fps to get enough frames to work from. This latency of 181-215ms is substantially worse than the 60ms latency between head movement & resultant VR changes being displayed that is often quoted as the upper limit for an acceptable VR experience.

The HMD framerate throughout the sessions varied between 19 and 60 fps. The median value was 48 fps and apart from a few outliers

[5] Walking round St Salvators https://www.youtube.com/watch?v=W4oPIHIr9Z4 Three views walking round St Salvators: https://www.youtube.com/watch?v=pvGV5dCjt4U

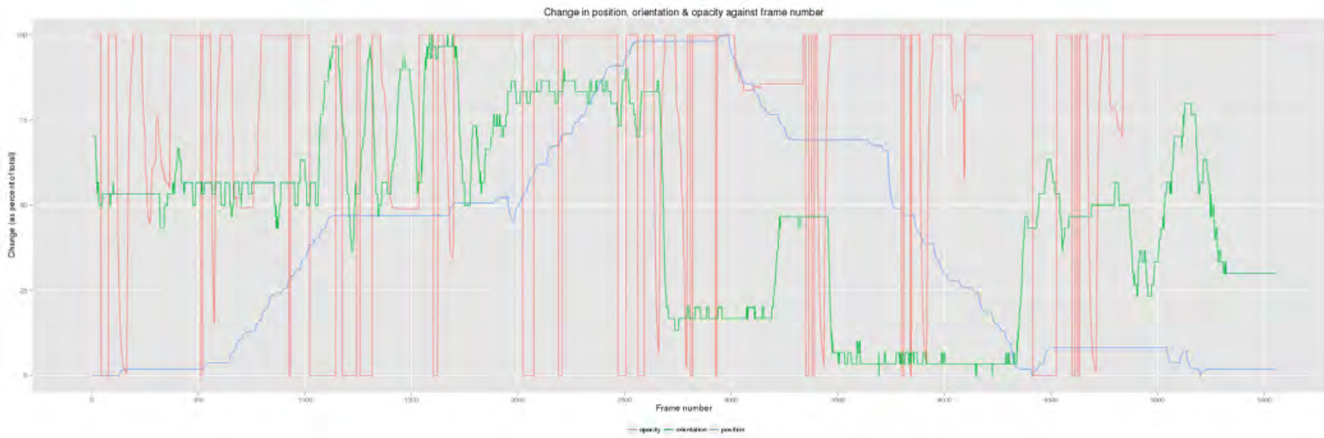

Figure 24: *Plot of a subset of example objective quantitative data logged by the Unity application.*

a framerate of above 30 fps was achieved. This proved to be satisfactory for this use case. Interestingly, whilst motion sickness had been widely reported in using the Oculus Rift whilst navigating with a key board or game controller. It was not reported during these experiments.

7 User Evaluation

Figure 24 shows three of the fields of the automatic data logging, <left_opacity>, <left_rotation> & the sum of <delta_x> & <delta_z>, plotted against <frame_number> for a set of data collected during a preliminary test. A precursory glance at this plot reveals that certain relationships can be gleaned between these series;

- Where the user is standing still (no change in the y axis for position, blue line) there is more marked head movement (more change in the y axis for orientation, green line); for example, frames 1100-1650, 2550-2950 & 4500-5000.
- Where the user is moving at a steady speed (steady climb or fall in y of position, blue line) we see the user either using the real world feed or switching to a semi opaque virtual feed. Towards the end of the session there is less use of the variable opacity control and more use of complete transitions to 100% virtual or 100% real. Complete transitions are more common when standing still; for example, frames 500-700, 3000-3300 & 4000-4100.

The feedback from users of the experience was very strong and compared favorably to experience in exploring the same model using a high resolution monitor keyboard and mouse. In particular the ability to easily switch between views enabled comparison between the two realities. A strong sense of immersion and engagement was achieved. This is despite the lower framerates and lower resolution of the HMD compared to a traditional monitor. The course nature of the movement tracking was not seen as being important. This is in part due to the self consistency of the two realities meaning precise alignment of views is less important than in augmented reality systems.

8 Conclusion

This paper has outlined the development of a cross reality system based upon the Oculus Rift and evaluated its use in enabling comparison of St Salvator's chapel as it is today and as it was in the 15th Century. The system performance measurements showed that framerates of between 40 and 40 fps were achieved, that latency from cameras was around 180ms and that user position was tracked to within a few meters whilst moving and to within a meter whilst stationary.

An X-Box controller enabled users to switch between realities, by pushing a button or puling a trigger. There was a preference for alternating between real and virtual rather than viewing both simultaneously. The virtual was viewed more while stationary and the real was moving. The combination of easy switching and intuitive navigation effectively addressed the vacancy problem easy comparison between the two realities.

The strength of the immersive experience provided by stereoscopic vision, compensated for the low specs, in terms of framerate, resolution and accuracy of movement tracking provided by the system. All users found the experience to be extremely positive, enjoyable and informative.

References

ANDERSEN, P. B. G. Tangible Objects : Connecting Informational and Physical space.

CABRAL, M., ZUFFO, M., GHIROTTI, S., BELLOC, O., NOMURA, L., NAGAMURA, M., ANDRADE, F., FARIA, R., AND FERRAZ, L. 2007. An experience using X3D for virtual cultural heritage. In *Proceedings of the twelfth international conference on 3D web technology*, ACM, New York, NY, USA, Web3D '07, 161–164.

CHRISTOU, C., ANGUS, C., LOSCOS, C., DETTORI, A., AND ROUSSOU, M. 2006. A versatile large-scale multimodal VR system for cultural heritage visualization. In *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '06*, ACM Press, New York, New York, USA, 133.

DE AMICIS, R., GIRARDI, G., ANDREOLLI, M., CONTI, G., AND AMICIS, R. D. 2009. Game based technology to enhance the learning of history and cultural heritage. In *Proceedings of the International Conference on Advances in Computer Enterntainment Technology*, ACM, New York, NY, USA, ACE '09, 451.

DUBLON, G., PARDUE, L. S., MAYTON, B., SWARTZ, N., JOLIAT, N., HURST, P., AND PARADISO, J. A., 2011. DoppelLab:

Tools for exploring and harnessing multimodal sensor network data.

GELERNTER, D. 1993. *Mirror Worlds: or the Day Software Puts the Universe in a Shoebox...How It Will Happen and What It Will Mean*.

HEETER, C., 2003. Reflections on Real Presence by a Virtual Person.

IJSSELSTEIJN, W., DE RIDDER, H., HAMBERG, R., BOUWHUIS, D., AND FREEMAN, J., 1998. Perceived depth and the feeling of presence in 3DTV.

IJSSELSTEIJN, W. 2001. Understanding Presence. In *Proceedings of the AIIA 2002 'Workshop sulla percezione della presenza in ambienti virtuali o remoti'*.

KIM, K., SEO, B.-K., HAN, J.-H., AND PARK, J.-I. 2009. Augmented reality tour system for immersive experience of cultural heritage. In *Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry - VRCAI '09*, ACM Press, New York, New York, USA, 323.

LAUREYS, S., BOLY, M., MOONEN, G., AND MAQUET, P. 2009. Two Dimensions of Consciousness: Arousal and Awareness. 1133–1142.

LIFTON, J. 2007. *Dual Reality: An Emerging Medium*. Ph.d. dissertation, Massachusetts Institute of Technology, Department of Media Arts and Sciences.

MAGNENAT-THALMANN, N., PAPAGIANNAKIS, G., AND CHAUDHURI, P. 2008. Interactive virtual humans in mobile augmented reality. In *Encyclopedia of Multimedia*. 362–368.

MAGNENAT-THALMANN, N., FONI, A. E., PAPAGIANNAKIS, G., AND CADI-YAZLI, N. 7. Real Time Animation and Illumination in Ancient Roman Sites. *The International Journal of Virtual Reality2 6*, 1, 11–24.

MANN, S., AND FUNG, J., 2002. EyeTap Devices for Augmented, Deliberately Diminished, or Otherwise Altered Visual Perception of Rigid Planar Patches of Real-World Scenes.

MILGRAM, P., AND KISHINO, F. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems E77-D*, 12, 1321–1329.

OKURA, F., KANBARA, M., AND YOKOYA, N. 2006. Fly-through Heijo Palace Site : Augmented Telepresence Using Aerial Omnidirectional Videos. *Computer 1*, 4503–4503.

OKURA, F., KANBARA, M., AND YOKOYA, N. 2011. Fly-through Heijo Palace Site: Historical Tourism System Using Augmented Telepresence. In *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2011, Vancouver, BC, Canada, August 7-11, 2011 , Poster Proceedings*.

PAPAGIANNAKIS, G., AND MAGNENAT-THALMANN, N. 2007. Mobile Augmented Heritage: Enabling Human Life in Ancient Pompeii. *International Journal of Architectural Computing 5*, 2, 396–415.

PAPAGIANNAKIS, G., AND MAGNENAT-THALMANN, N. 2009. Recreating Daily Life in Pompeii. *Arqueologica 2.0 1*, 16–20.

PAPAGIANNAKIS, G., SCHERTENLEIB, S., PONDER, M., ARÉVALO-POIZAT, M., MAGNENAT-THALMANN, N., AND THALMANN, D. 2004. Real-time virtual humans in AR sites. In *1st European Conference on Visual Media Production CVMP*, IET, no. 4, 273–276.

PAPAGIANNAKIS, G., SCHERTENLEIB, S., O'KENNEDY, B., AREVALO-POIZAT, M., MAGNENAT-THALMANN, N., STODDART, A., AND THALMANN, D. 2005. Mixing virtual and real scenes in the site of ancient Pompeii. *Computer Animation And Virtual Worlds 16*, 1, 11–24.

QVORTRUP, L. 2002. Cyberspace as Representation of Space Experience: In Defence of a Phenomenological Approach. In *Virtual Space: Spatiality in Virtual Inhabited 3D Worlds*, L. Qvortrup, Ed. Springer-Verlag, ch. 1, 5–24.

ROUSSOU, M., DRETTAKIS, G., AND MARIA ROUSSOU, G. D. 2003. Photorealism and Non-Photorealism in Virtual Heritage Representation. In *Proceedings of the International Symposium on Virtual Reality, Archeology and Cultural Heritage (VAST)*.

ROUSSOU, M. 2002. Virtual Heritage : From the Research Lab to the Broad Public. In *VAST 2000 Euroconference*, BAR International Series 1075, Oxford, Archaeopress, 93–100.

SLATER, M., AND STEED, A. 2000. A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments 9*, 5 (Oct.), 413–434.

SLATER, M. 2002. Presence and The Sixth Sense. *Presence: Teleoperators and Virtual Environments 11*, 4 (Aug.), 435–439.

TAKETOMI, T., SATO, T., AND YOKOYA, N. 2011. AR Cultural Heritage Reconstruction Based on Feature Landmark Database Constructed by Using Omnidirectional Range Sensor. In *Computer Vision ACCV 2010 Workshops*. 265–275.

TURKLE, S. 2004. Our Split Screens. In *Community in the Digital Age: Philosophy and Practice*, A. Feenberg and D. Barney, Eds. Rowman & Littlefield, ch. 6, 101–117.

TZORTZAKI, D. 2002. Virtual Reality as Simulation: The CAVE as "Space of Illusion" in Museum Displays. In *Virtual Space: Spatiality in Virtual Inhabited 3D Worlds*, L. Qvortrup, Ed. Springer-Verlag, 258–284.

VLAHAKIS, V., IOANNIDIS, N., KARIGIANNIS, J., TSOTROS, M., GOUNARIS, M., ALMEIDA, L., STRICKER, D., GLEUE, T., CHRISTOU, I. T., AND CARLUCCI, R. 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*, ACM, New York, NY, USA, VAST '01, 131–140.

WALCZAK, K., AND WHITE, M. 2003. Cultural heritage applications of virtual reality. In *Proceedings of the eighth international conference on 3D Web technology*, ACM, New York, NY, USA, Web3D '03, 182–183.

WATERWORTH, E. L., AND WATERWORTH, J. A. 2001. Focus, locus, and sensus: the three dimensions of virtual experience. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society 4*, 2, 203–213.

WILLMOTT, J., WRIGHT, L. I., ARNOLD, D. B., AND DAY, A. M. 2001. Rendering of large and complex urban environments for real time heritage reconstructions. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, ACM, New York, NY, USA, VAST '01, 111–120.

WOJCIECHOWSKI, R., WALCZAK, K., WHITE, M., AND CELLARY, W. 2004. Building Virtual and Augmented Reality museum exhibitions. In *Proceedings of the ninth international conference on 3D Web technology - Web3D '04*, ACM Press, New York, New York, USA, 135.