

Parallel Reality: Tandem Exploration of Real & Virtual Environments

CJ Davies



**University of
St Andrews**

This thesis is submitted in partial fulfilment for the degree of PhD
at the University of St Andrews

March 2015

Parallel Reality: Tandem Exploration of Real & Virtual Environments

CJ Davies

March 2015

1. Candidates declarations

I, CJ Davies, hereby certify that this thesis, which is approximately 27 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in [month, year] and as a candidate for the degree of PhD in [month, year]; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

2. Supervisors declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

3. Permission for publication

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

a. No embargo on print copy

Signature of supervisor

Abstract

- **Content** - Short, high-level summary of document.
 - **What has been done** - Nothing.
 - **What is left to do/how long should it take** - Write entire abstract, but should be quick to do & be done (near) last.
-

Acknowledgements

- **Content** - Acknowledgements.
 - **What has been done** - Nothing.
 - **What is left to do/how long should it take** - Short list of acknowledgements, should be quick & one of the last things to do, not overly important.
-

Contents

1	Introduction	6
2	Extended Example	7
3	Background, Theory & Rationale	8
3.1	Introduction	9
3.2	Defining Alternate Realities	10
3.2.1	Reality Continua	10
3.2.2	Reality Matrix	12
3.2.3	More Reality Continua	13
3.3	Adopted Definitions of Alternate Realities	13
3.3.1	Reality	14
3.3.2	Virtual Reality	14
3.3.3	Mixed Reality	15
3.3.4	Augmented Reality	15
3.3.5	Augmented Virtuality	15
3.4	Cross Reality	15
3.4.1	Overview of Cross Reality	16
3.4.2	MIT Media Lab Responsive Environments Group	16
3.4.3	Dual Reality: An Emerging Medium	17
3.4.4	IBM Virtual Universe Community	18
3.4.5	Shadow Lab	19
3.4.6	Ubiquitous Sensor Portals	20
3.4.7	Doppellab	21
3.4.8	Consumer Adoption of Cross Reality	22
3.4.9	X-Reality	23
3.4.10	IEEE Pervasive Computing - Cross Reality Environments	23
3.4.11	Eolus One	24
3.4.12	IBM and Implenia	25
3.5	Components of a Cross Reality System	25
3.5.1	Virtual Environments	25
3.5.2	Sensor/Actuator Infrastructure	27
3.5.3	Integration of Sensor/Actuator Infrastructure with Virtual Environments	31
3.6	Use Cases	34
3.6.1	Virtual Time Windows - Spatial Equivalence	34
3.6.2	Snow Crash - No Spatial Equivalence	35
3.7	Conclusions	36
4	Virtual Time Window	38

5 Mirrorshades - Design/Implementation	39
5.1 Introduction	39
5.1.1 Position - <i>Milgram & Kishino</i>	39
5.1.2 Position - <i>Waterworth & Waterworth</i>	42
5.1.3 The Case for Mobile XR	43
5.1.4 Example Application - Cultural Heritage	43
5.1.5 Case Study - St Salvator's Chapel	44
5.2 The Mirrorshades Platform	45
5.2.1 Implementation	45
5.2.2 Hardware Components	46
5.2.3 Software Components	46
5.2.4 Integration of Components	47
5.3 Investigation 1 - The Case for Mobile XR	47
5.3.1 Setting & Task	47
5.3.2 Evaluation Techniques	47
5.3.3 Hypothesis	48
5.4 Investigation 2 - Informing Mobile XR Implementation	48
5.4.1 Transitions using the Combined Model	49
5.4.2 Transition Methods	49
5.4.3 Experimental Task & Setting	54
5.4.4 Evaluation	54
5.4.5 Preliminary Results	56
5.5 IPQ Items	58
6 Mirrorshades - Studies/Results	59
7 Conclusions & Discussion	80
Appendices	81
Bibliography	81

1

Introduction

- **Content** - Short (10 pages is probably far too much). High level introduction to the concepts/topics involved in the thesis, very short/broad definitions of any terms introduced in the title (eg ‘simultaneous presence’), overview of the rest of the document, list of contributions/donations of this document
 - **What has been done** - Nothing.
 - **What is left to do/how long should it take** - Should be one of the last things to be written, maybe a few days of writing.
-

2

Extended Example

- **Content** - Short (10 pages is probably far too long) example or usage scenario of how the concepts investigated in the thesis have/could be used (a ‘near-future usage scenario’).
 - **What has been done** - Nothing.
 - **What is left to do/how long should it take** - Come up with a legitimate scenario, keep it short, should probably be written after the bulk of the later sections so there’s a clear idea of what scenario I actually want to allude to. Perhaps a few days writing.
-

3

Background, Theory & Rationale

- **Content** - Should lay out the theoretical motivations behind this work, outline the main issues that arise given these motivations & describe the approaches taken to overcome the issues. For each part of this, pertinent related work should be overviewed (eg. this is where the literature review will be situated).
- **What has been done** - Extensive literature review was written at the end of year one. This literature review needs re-writing such that the pertinent parts are included in the relevant part of the discussion of ‘theory & rationale’ (eg the content of the literature review will be split up & inserted where appropriate into the new ‘theory & rationale’ document). Literature read since the end of year one (which is substantial & includes a slight shift in focus; presence, PoSR, older VR & AR literature) needs incorporating.
- **What is left to do/how long should it take** - New document needs writing in the guise of ‘theory & rationale’, splitting up & inserting contents of old literature review & writing in about more recently read literature since. 30 pages of the current literature review document is far too long, needs to be substantially cut down whilst covering more literature. Substantial exercise in writing/collating/drawing links between, etc. Might want to put aside a month or so of write-up time for this alone.

“Where are you?” Hiro says.

“In Reality or the Metaverse?”

“Both.”

Snow Crash, Neal Stephenson

Reviewing the literature on the domain of alternate realities this research finds that there is a gap in the scholarly investigation of simultaneous presence in real and virtual environments and the associated ‘vacancy problem’. This review proposes that a better understanding of the extension of human presence from only one of the real world or a virtual environment to simultaneous presence in both will permit the introduction of novel systems in a variety of fields in which simultaneous interaction and exploration of both real and virtual environments is possible. Such systems will likely be formed by expansion of research into *cross reality*, an alternate reality comprised of complete real and virtual environments able to mutually reflect and influence each other via sensor/actuator infrastructure, and are likely to be in high demand as progress toward 3D extension of the Web continues.

3.1 Introduction

Research on alternate realities has been extensive, with the theme being explored for purposes as diverse as education [1] and new forms of data visualisation [2] to medical [3] and military training [4]. Traditionally access to implementations of alternate realities was limited by the availability and high monetary cost of the specialised hardware and software that was required to create virtual environments and computerised superimpositions [5]. However more recently the cost and availability of this hardware and software has reduced and increased respectively to the point at which we find ourselves today where alternate realities can be, and in fact frequently are, experienced by many with the equipment they already have available to them with no special considerations [6].

Hardware capable of synthesising the graphically complex three-dimensional environments that alternate realities often present has reduced in cost to the point that it is no longer considered specialist equipment, owned only by those expressing an explicit interest in the subject, and is now commonplace in ‘average’ consumer electronics devices, be they traditional desktop and laptop computers, games consoles, or portable devices such as mobile phones and tablets. Combined with the continued adoption of high-speed Internet connectivity [7] the potential of multi-user virtual environments is already being realised, both for traditional competitive gaming through platforms such as World of Warcraft, but also for non-competitive purposes that focus more on community, creation and commerce with ‘virtual world’ platforms such as Second Life [6].

Furthermore, progress toward ubiquity of sensing and actuating infrastructure in our everyday lives continues at an accelerated rate [8, 9]. Such systems are now commonplace in new buildings and also in consumer electronics, both portable devices such as mobile phones and tablets as well as home entertainment products such as games consoles. The result is that the amount of information that we can access about the physical and environmental state of a particular location at any given time is greater than ever.

The combination of these factors means that we are approaching the required dissemination of technology and public knowledge and understanding for virtual environments and integration with sensor/actuator infrastructure to begin making a larger impact on society and ultimately to be used by on a scale and in a manner similar to how the World Wide Web already is today. It is no longer preposterous to imagine a near future where a ‘Metaverse’ reminiscent of that in Neal Stephenson’s cyberpunk novel *Snow Crash*, presenting an extension of the Web in the form of a three-dimensional virtual environment, has become a reality adopted by the majority of current Web users rather than remaining a vision of academics and cyberpunk fiction aficionados.

However although there are numerous examples of virtual environments, particularly those for competitive gaming, gaining substantial popularity [6] and augmented reality and augmented virtuality products are no longer restrained to the research lab, this review has found very little research into simultaneous presence in *complete* real and virtual environments. The term *complete* is used to emphasise that the discussion here is about real and virtual environments that are both complete unto themselves and which can thus be explored and interacted with in isolation from the other, in addition to being explored simultaneously. This is a different concept to augmented reality and augmented virtuality systems, where the augmentations are usually near meaningless when separated from the context bestowed upon them by the underlying environment that they are augmented upon.

The *cross reality* paradigm represents the most promising foray into investigation of this concept, as such a system is established through the combination of two complete environments, a real environment and a virtual environment that is based upon and mimics this real environment, which are bestowed with the abilities of mutual reflectance and influence through the use of sensor/actuator infrastructure (see section 3.4 for full discussion). Physical and environmental changes in the state of the real environment are captured by sensors and these data used to update the state of the virtual environment, whilst simultaneously changes in the state of the virtual environment manifest into the real environment via actuators. Users are free to explore and interact with either environment in relative isolation from the other, even if their interactions in one trigger changes in the other, however simultaneous interaction and exploration with both environments has largely remained without systematic investigation.

This is largely because users exploring and interacting with the real environment do not have a convenient manner of also exploring and directly interacting with the virtual environment, as such interaction usually

relies upon the use of software run on a desktop or laptop computer which is not conducive to mobile use. Using a laptop computer whilst walking around is far from convenient and using a desktop computer obviously limits the user's interaction with the real environment to that immediately around the location of the computer and results in a disjoint relationship between their physical position in the real environment and the location of their avatar in the virtual environment when they navigate their avatar away from the respective position of their computer. This situation has been called 'the vacancy problem'; an apparent vacancy from one environment whilst engrossed in the other (see section 3.4.3 for full discussion).

Interested in the promise of simultaneous presence in complete real and virtual environments, particularly those that are able to mutually affect each other via sensor/actuator infrastructure and the cross reality paradigm, this literature review investigates how academic studies have treated such themes. The conclusion of this investigation is that the theme of simultaneous presence in real and virtual environments is still discussed only marginally in the scholarly literature and that this situation warrants attention as the benefits proposed by such interactions between real and virtual environments are extremely promising.

3.2 Defining Alternate Realities

A fundamental imperative for the remainder of this review is a well defined set of criteria for classifying and differentiating between different types of alternate reality. The terms *mixed reality*, *augmented reality* and to a lesser extent *augmented virtuality* are all now relatively common in the literature, however they have too often been used in conflicting manners, or assigned vague definitions with uncertain boundaries separating them from each other. Furthermore the less well established terms *cross reality* and *X reality* (different names for the same concept) also need classifying according to the same system.

3.2.1 Reality Continua

Paul Milgram, Herman Colquhon and Fumio Kishino addressed this issue in detail and can in fact be accredited with introducing the terms *augmented virtuality* and *mixed reality* to the literature in the first instance, prompted by their identification of the need for more encompassing terms to supplement the existing definitions of *augmented reality* [10, 11]. Their discussion at times takes on a hint of the philosophical, as it rightly discusses what exactly it is that we mean by 'real' and 'virtual' and whether it is in fact reality or virtuality which is being augmented. However despite these thorough and well-reasoned definitions being published originally in 1994, much of the subsequent literature studied for this review has adopted conflicting, or at least confusing and misleading, definitions.

One of the overbearing concepts that Milgram et al. introduced is that whilst both purely real and purely virtual environments do exist they should not be considered discrete alternatives but rather poles lying at opposite ends of a linear scale called the *Reality-Virtuality continuum*. The location of an environment along this continuum coincides with its location along a parallel *Extent of World Knowledge continuum* where 'world knowledge' refers to the amount of quantitative information that is associated with the content being presented, or in other words how much of the environment is being 'modelled' by a computer. These continua are included as figure 3.1.

With a purely virtual environment, the entire viewport must necessarily be computer modelled in order to be rendered and as such there is complete quantitative information about all objects and between all objects being presented. At the opposite end of the spectrum with a completely real environment where none of the viewport is computer modelled there is no quantitative information associated with the content being displayed. At any point between the extremes the environment consists of a mixture of some modelled and some non-modelled content; with the computer associating quantitative information to, and between, the virtual objects, but not to the real objects or between the virtual and real objects.

Carrying the continuum concept further, Milgram et al. illustrate their understanding of the existing term *augmented reality* and also introduce two related new terms; *augmented virtuality* and *mixed reality* (see figure 3.2). In this fashion, *mixed reality* is used to describe any environment that is not completely real or completely virtual; that is, it encompasses all positions on the continuum between the extremes. *Augmented*

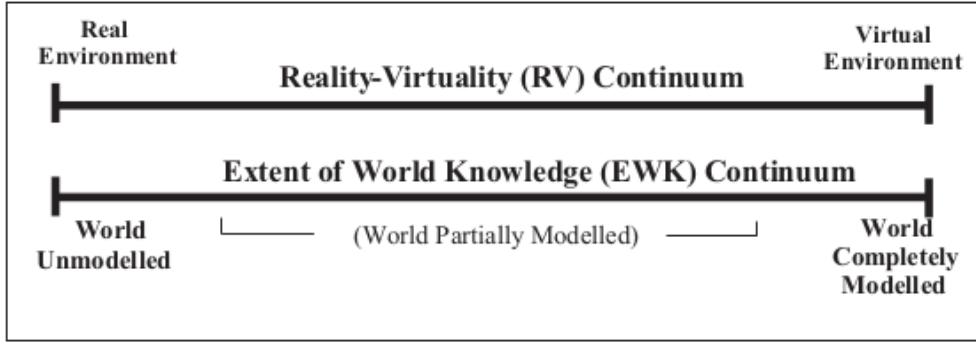


Figure 3.1: *Reality-Virtuality continuum*, parallel with *Extent of World Knowledge continuum*.

reality is used to describe a real environment upon which virtual objects are overlain and *augmented virtuality* is used to describe a virtual environment upon which objects sampled from the real world (such as video feeds) are overlain. It is also shown here that *mixed reality* encompasses both *augmented reality* and *augmented virtuality*.

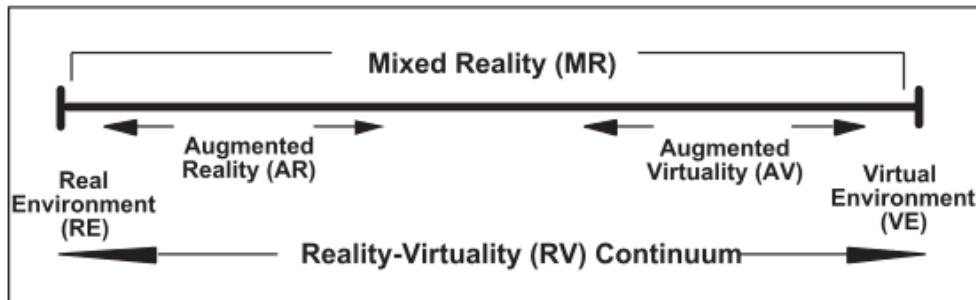


Figure 3.2: Illustration of the terms *mixed reality*, *augmented reality* and *augmented virtuality* within the context of the *Reality-Virtuality* continuum.

The obvious question raised from studying figure 3.2 is at what point toward the centre of the continuum an environment changes from being *augmented reality* into *augmented virtuality* or vice-versa. The answer lies with consideration of the quantitative knowledge associated with the objects that comprise the viewport.

For example, if one were to take a viewport depicting a purely real environment and then incrementally add more and more virtual objects, the environment's classification would progress rightward along the continuum. Eventually the entire viewport would be obscured by virtual objects and the obvious conclusion would be to classify the environment as being purely virtual. However this would only be true if there was complete quantitative information associated with, and between, all of the virtual objects within the real 3D space of the viewport, which is unlikely to be the case.

Likewise if one were to take a viewport depicting a purely virtual environment and incrementally replace the entire viewport with sampled real objects we could not classify the resultant environment as purely real as there would be associated quantitative knowledge with and between the sampled objects, meaning that the environment isn't completely unmodelled and thus can't be classified as purely real.

Thus, Milgram et al. conclude, it is not necessarily true that an environment is purely virtual simply because all of the visible objects are computer modelled, nor is it necessarily true that an environment is purely real simply because all of the visible objects are sampled from the real world.

3.2.2 Reality Matrix

Another useful method of illustrating the relationships between the different categories of alternate realities was put forward by Roy Want in his introductory article for a 2009 issue of IEEE Pervasive Computing dedicated to the *cross reality* paradigm [12]. Here he presented a 2x2 matrix categorising the different terms according to whether the experience and overlay data are real or virtual (see figure 3.3).



Figure 3.3: Want's original virtuality matrix.

Whilst this is a useful representation, some of the definitions and criteria depicted do not match with those of Milgram et al. or even with those of other authors in the same issue of IEEE Pervasive Computing, let alone other publications concerning alternate realities. Thus this review presents a modified version of Want's matrix, that is better in fitting with the consensual definitions built from reviewing the literature on alternate realities, as figure 3.4.

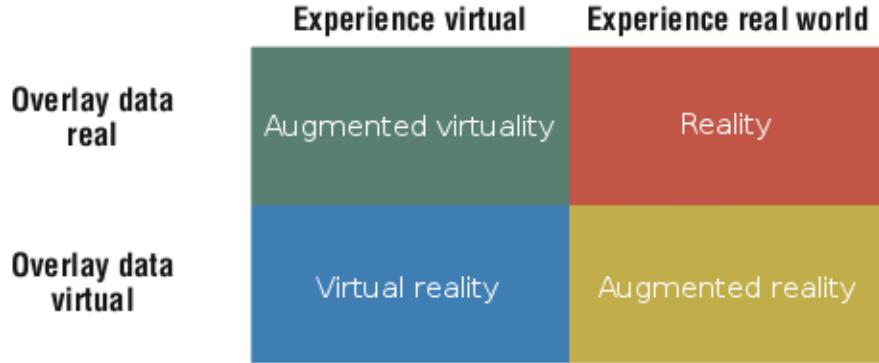


Figure 3.4: Want's virtuality matrix after modification by this review; note removal of *mixed reality*, *cross reality* and *embodied virtuality* and addition of *reality* and *augmented virtuality*.

Where Want has *cross reality* occupying the upper left quadrant at the congruence of 'experience virtual' and 'overlay data real' this review instead places *augmented virtuality*. Referring to Milgram's continuum 'experience virtual' relates to a position somewhere on the right half, while 'overlay data real' relates to presentation over this virtual environment of sampled real world data resulting in a partially modelled

environment, leaving us in the area of the continuum occupied by *augmented virtuality*.

Want's matrix also features the term *embodied virtuality* in the upper right quadrant, at the congruence of 'experience real world' and 'overlay data real'. Want explains that this is an alternative term for *ubiquitous computing* which is "*essentially the opposite of VR*"; this review instead reasons that the opposite of *virtual reality* is simply *reality* and that *ubiquitous computing* does not constitute an alternate reality but simply a different model of human-computer interaction that can be implemented in either *reality* or *augmented reality*, depending upon how the computing infrastructure presents information to users.

A *ubiquitous computing* system is necessarily a real environment, as it is by definition the integration and dissemination of computational infrastructure into our real surrounds [13]. However whether this real environment is augmented by virtual objects is not restricted by the concept. Thus *ubiquitous computing* can exist in an environment that is either on the left extreme of the continuum, where no virtual objects are employed and the environment remains completely unmodelled, or somewhere to the right of the left extreme in the region of *augmented reality*, where virtual objects are employed and the environment is partially modelled.

This line of reasoning is supported by an almost complete lack of further mention of *ubiquitous computing* elsewhere in the literature about alternate realities studied for this review. Furthermore the term *embodied virtuality* is not used by any other author in the studied literature.

Finally this review has removed the central *mixed reality* section from Want's original matrix, as its position could be misleading. As the boundaries formed between the categories by the different colours could be construed as meaning that there are discrete boundaries between the different categories, rather than a linear scale as depicted by Milgram's continuum, the reader could be led to believe that a purely *virtual reality* or a purely *embodied virtuality* environment can be considered *mixed reality*, which is incorrect. If one wishes to picture the position of *mixed reality* in relation to the modified matrix, it would cover the same area as enclosed by the union of *augmented virtuality* and *augmented reality*.

3.2.3 More Reality Continua

As one of the most prominent academics in the development of the cross reality paradigm (see section 3.4) it is also worth comparing Joshua Lifton's definitions for the different categories of alternate realities and the relationships between them [14] in the current context.

Lifton's definitions, or more accurately his relationships between, the terms *reality*, *augmented reality*, *mixed reality* and *virtual reality* do not perfectly match the consensus that this review has observed. Lifton defines the terms individually in agreement with the consensus, however doesn't proffer the conclusion that mixed reality is a broad term that includes augmented reality. He also doesn't mention augmented virtuality, even though the Dual Reality Lab project (see section 3.4.5) does implement it. Finally his diagram that alludes to Milgram's continua and is included as figure 3.5, situates mixed reality at an incorrect position, implying that Lifton's definition of mixed reality is of a discrete state to that of augmented reality, even though his textual definition of mixed reality hints that it logically encompasses augmented reality. This review presents a modified version of Lifton's diagram to illustrate these differences as figure 3.6.

Lifton does however explain that while such a taxonomy can be successfully applied to most alternate reality efforts, it does not well address the concept of cross reality where there are two complete realities, one real and one virtual.

3.3 Adopted Definitions of Alternate Realities

This review has discovered differing (and in some cases conflicting) definitions for the different categories of alternate realities and for the criteria for differentiating between them. What follows in this section represents the definitions and differentiating criteria that this review has adopted after concluding them the most widely accepted and well reasoned.

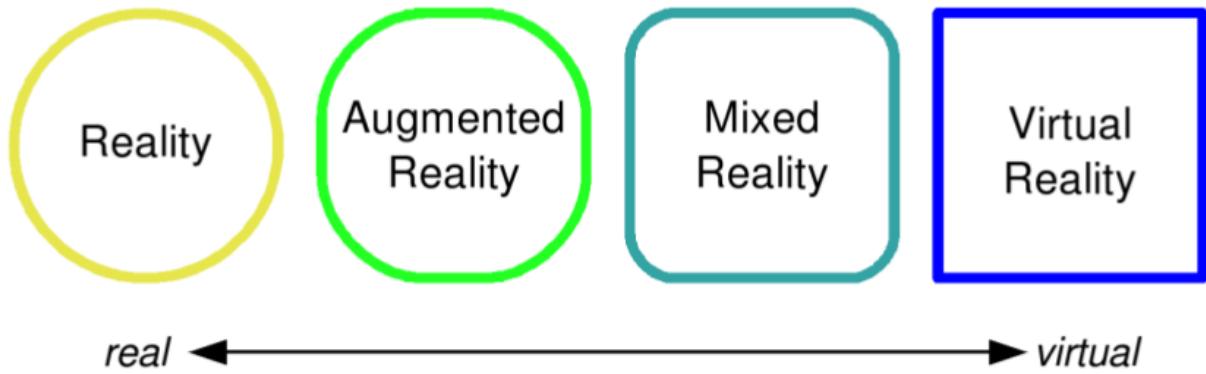


Figure 3.5: The “*virtual worlds taxonomy as viewed on the real-virtual axis*” presented by Lifton.

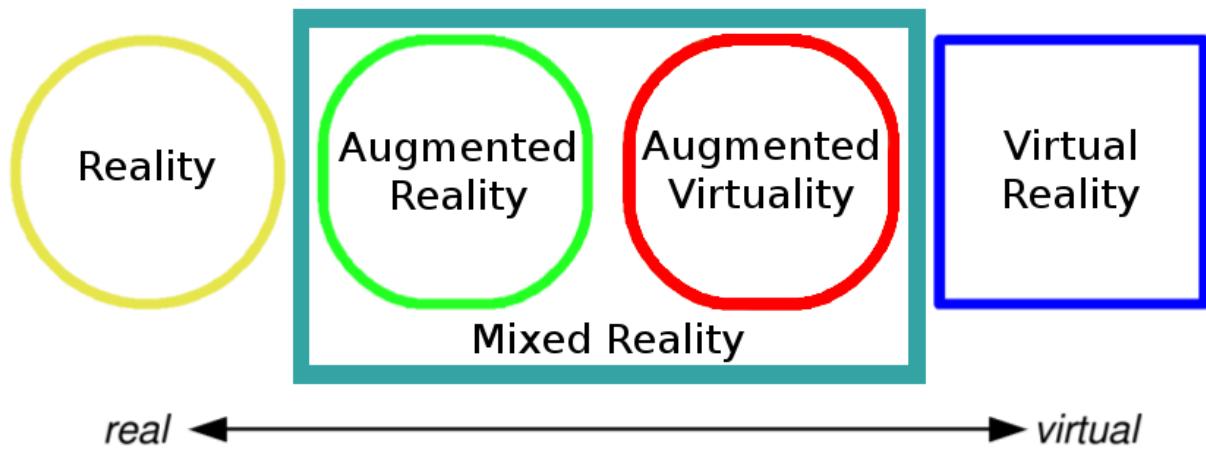


Figure 3.6: Lifton’s taxonomy as modified by this review.

3.3.1 Reality

Occupying the left extreme of Milgram’s continuum and the upper right quadrant of the modified Want matrix, *reality* refers to an environment that is entirely unmodelled, with the viewport containing no virtual objects and no computer-based quantitative information associated with any of the (necessarily real) objects. In fact, there may be no computer infrastructure involved in the situation whatsoever. This is the situation that we are most familiar with, as it is where the vast majority of us spend the vast majority of our time.

3.3.2 Virtual Reality

The polar opposite to *reality*, a *virtual reality* environment occupies the right extreme of Milgram’s continuum and the lower left quadrant of the modified Want matrix. A *virtual reality* environment consists solely of virtual objects, with computer-based quantitative information associated with all of them and between all of them, creating a completely synthetic world entirely discrete and separate from the real world; a new world that exists solely within the data structures of a computer.

Traditional definitions of *virtual reality* require the environment to be completely immersing; that is, when involved with such an environment the user is completely unaware of the real environment that sur-

rounds their real bodies, often making use of Head Mounted Displays (HMD) and head &/or body tracking techniques to improve the sense of immersion by removing the need to interact with interfaces logically anchored to the real world such as keyboards, mice and joysticks [5].

However this review believes that taking the concept to a lesser extreme, the virtual environments presented by video games such as World of Warcraft can be construed as *rudimentary* implementations of virtual reality; they are completely modelled environments that exist entirely separate to the real world, however interaction is not completely immersing due to the use of 2D monitors and traditional interface devices, largely due to the lack of common ownership of HMDs and advanced body tracking systems.

3.3.3 Mixed Reality

Occupying any position between the extremes on Milgram's continuum, the term *mixed reality* refers to a broad range of environments that arise from the merging of real and virtual environments to some extent such that the result is neither entirely real nor entirely synthetic, where real and virtual objects co-exist. As alluded to previously, under this definition both *augmented reality* and *augmented virtuality* are included under the broader classification of *mixed reality*.

This is one definition where Want is lacking, claiming *mixed reality* to be “*some combination of the others*” where ‘others’ refers to all of *virtual reality*, *augmented reality*, *embodied virtuality* and *cross reality*. This review disregards this definition as it relies upon Want’s previously refuted definition of *embodied virtuality* being the polar opposite of *virtual reality* and because it also depends upon Want’s definition of *cross reality* which this review will also go on to refute.

3.3.4 Augmented Reality

An *augmented reality* environment occupies a position within the ‘left half’ of Milgram’s continuum, the lower right quadrant of the modified Want matrix and within the broader classification of *mixed reality*. Thus an *augmented reality* environment comprises a real environment that has had virtual objects added to or overlain upon it; a common approach for achieving this addition/overlay is superimposing virtual objects over a direct or indirect view of the real environment using Head Mounted Displays &/or cameras [15].

A commercial example of *augmented reality* is the Layar browser for mobile phones, which overlays various forms of data onto the view captured by a phone’s camera after determining its location and orientation using GPS, accelerometer and magnetometer readings [16]

3.3.5 Augmented Virtuality

Logically opposite to *augmented reality*, an *augmented virtuality* environment occupies a position within the ‘right half’ of Milgram’s continuum, the upper left quadrant of the modified Want matrix and again lies within the broader classification of *mixed reality*. Thus an *augmented virtuality* environment comprises a virtual environment, akin to *virtual reality*, upon which sampled real objects are overlain, perhaps through the use of cameras [17].

A simple commercial example of augmented virtuality is the EyeToy accessory and associated software for Sony’s Playstation 2 games console (and later the Playstation Eye for the Playstation 3), a digital camera that captures images of players and their surroundings and integrates them into the gaming experience presented on the screen.

3.4 Cross Reality

The discussion of the literature pertaining to the *cross reality* paradigm warrants its own section of this review, as it is not only one of the youngest categories included under the umbrella term *alternate reality* but also the closest existing concept to lend to the simultaneous exploration of real and virtual environments.

First a brief overview of the paradigm is presented, that represents the conclusions drawn after studying the available literature on the subject. This is followed by an in-depth study of the literature that led to these conclusions, arranged largely chronologically to show what parties were responsible for the paradigm's creation and development from its inception.

As cross reality is a relatively recent concept that hasn't received a vast amount of scholarly attention, this section of the review represents an almost exhaustive survey of the literature published on the subject. For such a recent concept that is still open to debate and investigation, it is prudent to discuss all of the work that has thus far contributed to the field in any manner, rather than trying to judge contributions at such an early stage when the bounds of the subject matter are not yet well established or agreed.

3.4.1 Overview of Cross Reality

Cross reality is the ubiquitous mixed reality situation that arises from the fusion of real-world sensor/actuator infrastructure with virtual environments, such that augmented reality and augmented virtuality manifest simultaneously and facilitate synchronous multi-directional exchange of media and control information between real and virtual environments. Sensors collect and tunnel dense real-world data into virtual environments where they are interpreted and displayed to dispersed users, whilst interaction of virtual participants simultaneously incarnates into the real world through a plenitude of diverse displays and actuators [18].

The principle features that distinguish cross reality from other alternate realities that this review has thus far covered are;

1. a shift from single- to bi-directional information flow between real and virtual environments [19]
2. that both environments are complete unto themselves (but are enriched by their ability to mutually reflect, influence and merge into one another). [20]

Because a cross reality system is comprised of two complete environments, it cannot be placed at a single point on Milgram's continuum but instead occupies two separate points, one to the left and the other to the right of the centre. Similarly with the modified Want matrix, a cross reality system would occupy one location in the upper left quadrant of augmented virtuality and a second in the lower right quadrant of augmented reality.

3.4.2 MIT Media Lab Responsive Environments Group

The Responsive Environments Group at MIT's Media Lab, under the direction of Joseph Paradiso, deserve the accolade for the inaugural work in establishing cross reality as a field of research. In their own words, the group "*explores how sensor networks augment and mediate human experience, interaction and perception*" [21] and since 1995 they have produced prolific research in the domain of sensor architecture and wireless sensor clusters [22–32] and perceptive spaces [33–36], establishing a prime research environment, both in terms of experience and available technologies, for a concept such as cross reality to emerge.

The path to cross reality began with the 'Plug' project that created a sensor network comprising power strips imbued with sensing, computational and communicative abilities [37]. By basing the nodes on power strips the platform was ideally suited for broad and unobtrusive deployment in environments where people work and live as power strips already exist in such environments with ubiquity. A labelled photograph of a single Plug node is included as figure 3.7. In addition to sporting a number of different sensors, each Plug node was also able to individually control the output of the four 120Vac sockets, allowing the nodes to act as both input (sensing) and output (actuating) devices.

Parallel to the development of the Plug platform the group also developed two mobile platforms for interacting with the network of nodes. The Tricorder project [38], based around a Nokia 770 Internet tablet, presented users with a two-dimensional map centered and oriented about the Tricorder's physical position and allowed "*real-time point-and-browse*" functionality to browse the sensor data from nodes within the vicinity. The Ubicorder [39] project made use of a tablet laptop computer and allowed users to graphically define,

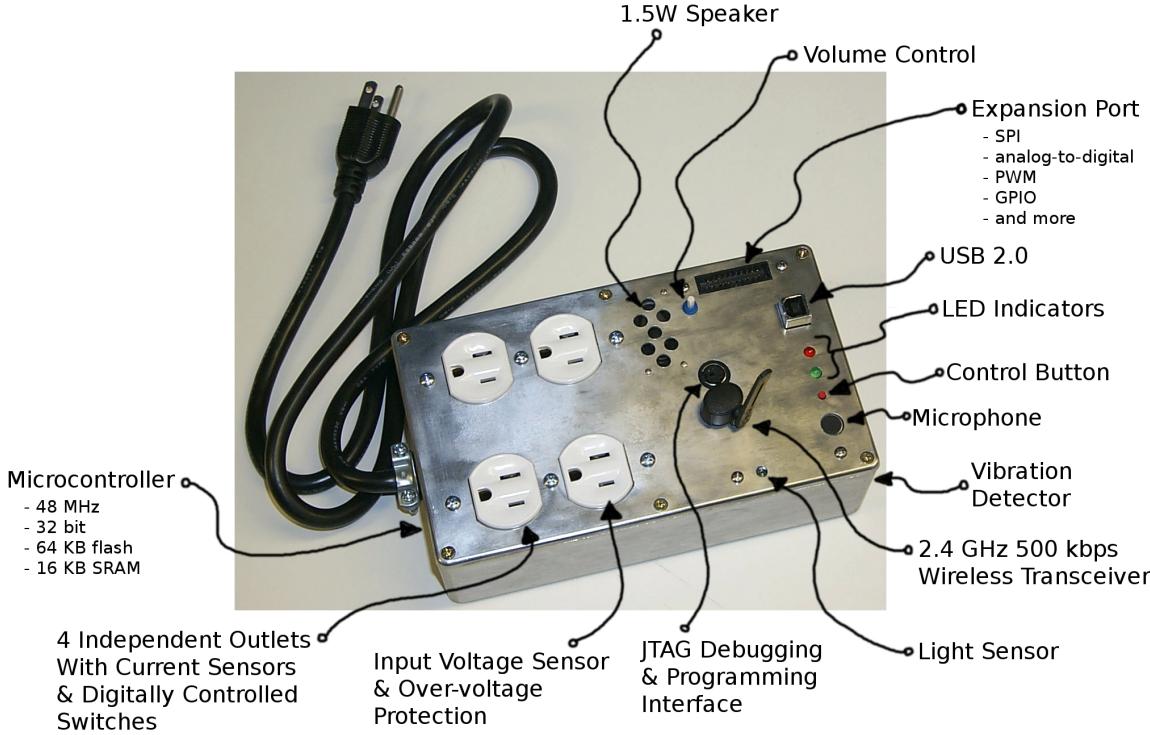


Figure 3.7: A Plug node.

and recursively combine, rules for translating sensor data to higher order and potentially more meaningful events.

The true birth of cross reality at the Media Lab came from the combination of the Plug platform with the Second Life virtual world in the Dual Reality Lab project which is comprehensively covered in Joshua Lifton's PhD thesis entitled 'Dual Reality: An Emerging Medium' [14] and summarized in [20]. The term 'dual reality' would later evolve into what we now know as 'cross reality', an association confirmed by Lifton himself [40]. As the inaugural work in establishing the field and framing cross reality as a new application domain for sensor/actuator infrastructure, virtual environments and media creation, and still by far the most exhaustive discussion of the paradigm, Lifton's thesis deserves closer attention in this review than other publications.

3.4.3 Dual Reality: An Emerging Medium

In his 2007 thesis Lifton defines dual reality as

"An environment resulting from the interplay between the real world and the virtual world, as mediated by networks of sensors and actuators. While both worlds are complete unto themselves, they are also enriched by their ability to mutually reflect, influence, and merge into one another." [14]

The key points to identify in this definition, which are confirmed later in the thesis and by other authors in subsequent literature, are that data flow is bi-directional between real and virtual environments and that the two environments are complete unto themselves in addition to having the ability to affect each other.

Perhaps the most interesting and topical discussion within Lifton's thesis for the purposes of this literature review, whose topic is simultaneous presence in complete real and virtual environments, is the discussion of what he calls '*the vacancy problem*', which he defines as

"the noticeable and profound absence of a person from one world, either real or virtual, while they are participating in the other. Simply put, the vacancy problem arises because people do not currently have the means to be in more than one place (reality) at a time. In the real world, the vacancy problem takes the form of people appearing completely absorbed in their virtual reality, to the exclusion of everything in the real world. In the virtual world, the vacancy problem takes the form of virtual metropolises appearing nearly empty because there are not enough avatars to fill them."

Lifton identifies that this vacancy problem is a fundamental characteristic of the current generation of virtual worlds and proposes dual reality, more closely linking the real world with the virtual world, as an approach to mitigate the problem.

Lifton also surveys a range of interesting previous work broadly related to the concept of coupling the real world with virtual worlds. He roughly categorises these into those that attempted to bring a virtual world into the real world (augmented reality) and those that attempted to bring the real world into a virtual world (augmented virtuality). However upon further investigation one of the projects Lifton mentions, from IBM UK, reveals experiments with two way control mechanisms between real and virtual worlds which alludes more directly to the cross reality paradigm.

3.4.4 IBM Virtual Universe Community

IBM was influential in the 2006-2009 wave of virtual worlds, their involvement starting through grass roots interests on internal blogs and on Second Life [41-43] from people such as Ian Hughes, at the time an IBM Software Strategist, and eventually expanding to include around 8000 employees (including the CEO at the time) and the creation of an Emerging Business Unit (EBO) - IBM's way to venture capital new ideas and see how they fit. A number of impressive projects were undertaken, the most high profile of which were the Wimbledon projects of 2006, 2007 and 2008 [44, 45]. Journalist Rita King provided a good write-up of the entire story of the 'Virtual Universe Community' at IBM [46] and various projects were featured by both Businessweek [47] and the BBC [48].

Ian Hughes confirmed via email to this review the extent toward a cross reality system that these investigations progressed

"The control mechanisms worked two ways generally. There was a physical lab that had devices that were controlled by a pub/sub mechanism based on the light weight protocol MQTT. Those devices subscribed to various messages. So initially web pages controlled them. The web page generated the message that was broadcast to everything that was interested, sending on/off messages. Equally the objects generated messages when they were physically switched on and off. As SL had an RPC interface it was possible (using another software component outside of SL) to subscribe to the same messages and send requests into SL to change states of object. Likewise it was simple to make an object when clicked or some other event in SL send a message back out. So there were lights, blinds, proximity detectors and even the tilt sensors on the laptops that were instrumented with these messages."

So whilst Lifton and the other researchers at the Media Lab cannot necessarily be credited with 'inventing' cross reality or performing the very first implementation of it, they deserve the accolade for being the first to perform an in-depth academic investigation into the concept, framing it as a new area of research interest.

3.4.5 Shadow Lab

The most impressive creation of the MIT Dual Reality Lab project was the ‘Shadow Lab’, “*a space in Second Life modelled after the third floor of the MIT Media Lab’s Wiesner building, where the Plug sensor network is most often deployed*”. This comprised a to-scale two-dimensional floor plan of the entire third floor of the building and a three-dimensional reconstruction of the Media Lab itself.

Around the Shadow Lab were distributed a number of ‘data ponds’ that each represented the data from a single Plug node. Each pond consisted of a number of waving stalks growing out of a puddle of water on the ground, with the different types of data affecting the appearance of different features; light affecting the pattern of the stalks’ skin, temperature affecting their colour, motion affecting the stalks’ motion, sound affecting the size of the puddle and electrical current (of devices drawing power from the 120Vac sockets of the Plug node) affecting the intensity of ethereal foxfire rising from amongst the stalks. These virtual data ponds allowed for the real-to-virtual, augmented virtuality aspect of the system. Inversely real world versions of the data ponds, each comprising a desk fan shrouded with lightweight plastic, allowed for the virtual-to-real, augmented reality aspect of the system. The airflow through the shroud, and thus the height and sound produced, could be controlled by pulse width modulating the output of the 120Vac sockets of the Plug nodes. Figure 3.8 shows a virtual data pond (at left) and a number of real data ponds (at right) and figure 3.9 shows the final Shadow Lab itself.

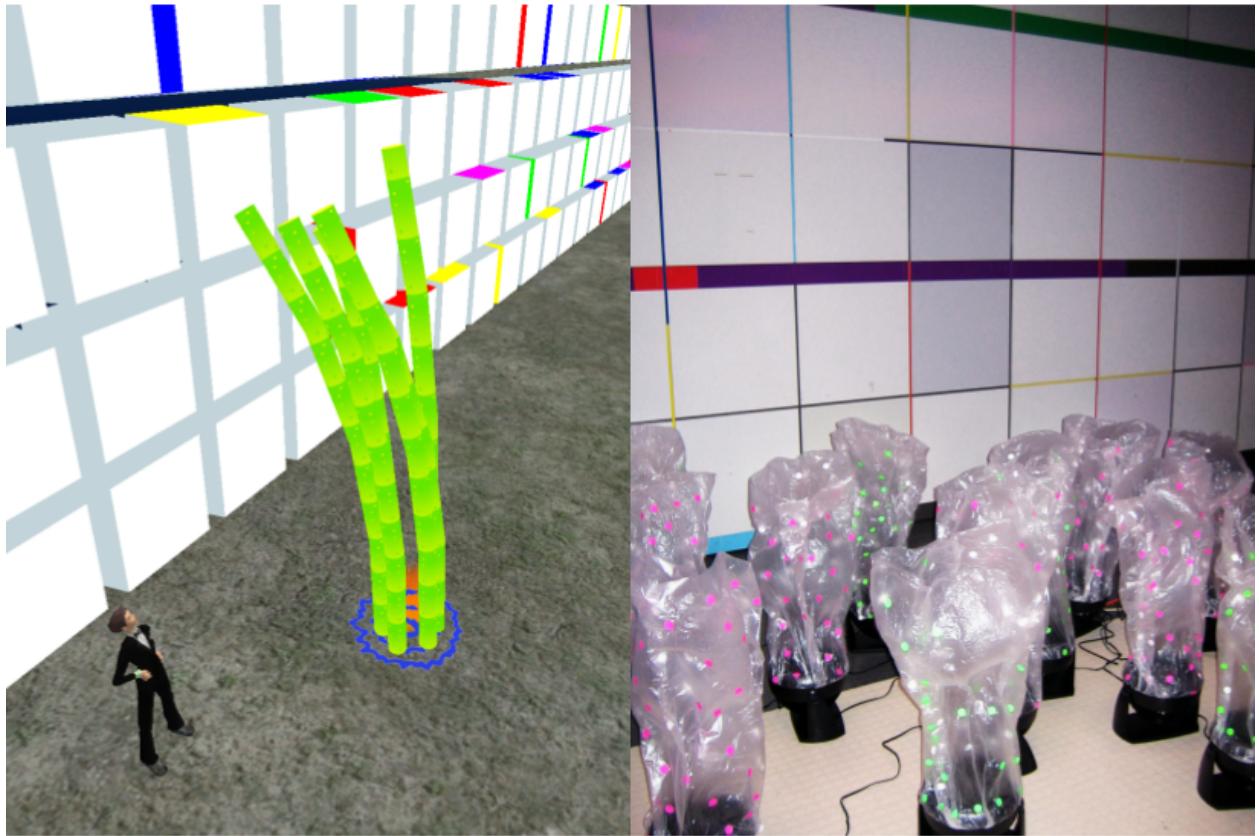


Figure 3.8: Virtual (left) and real (right) data ponds.

Lifton evaluates the Dual Reality Lab project and makes the prudent observation that widespread adoption of the cross reality paradigm will rely upon sensor/actuator infrastructure as pervasive as today’s electrical and lighting systems and online virtual worlds as populated as today’s World Wide Web, both of which are years if not decades away. He also notes that even if these technologies do come to pass, the



Figure 3.9: Side view of the final Shadow Lab structure, showing the two-dimensional floor plan of the third floor of the Wiesner building, three-dimensional reconstruction of the Media Lab, virtual data ponds and a human-sized avatar.

emergence of cross reality is not certain.

3.4.6 Ubiquitous Sensor Portals

Following on from the Dual Reality Lab, and now dropping the title dual reality in favour of cross reality, the Ubiquitous Sensor Portal project situated 45 I/O rich ‘portals’, shown in figure 3.10, throughout the Media Lab, each with a corresponding extension in Second Life. Each portal comprised a myriad of environmental sensors (passive infrared (PIR) motion, light, sound, vibration, temperature and humidity) in addition to multimedia abilities via a camera, mounted on a motorised pan/tilt platform and capable of capturing still images as well as DVD-quality video, microphone and small touch screen display. They also featured active IR links that allowed them to communicate with various wearable badges in development by Media Lab researchers to identify users stood at a portal, and in addition could be used as reflection proximity sensors to detect the presence of unbadged users. When a user is present at a real portal, a white ‘ghost’ is displayed by the corresponding virtual portal if they cannot be identified by badge, and their name displayed if they can. Video and audio from the real portal can be streamed into Second Life and a stream from a Second Life client streamed to the screen of the real portal.

However in stark contrast to the Dual Reality Lab, the virtual portals were not situated in a simulation of the real Media Lab in situations corresponding to their physical location, but instead used a more ab-



Figure 3.10: A Ubiquitous Sensor Portal.

stract virtual representation with a geometric layout; this design, shown in figure 3.11, reflected intellectual affiliation as opposed to real-world location and also allowed for intuitive browsing of past still images and videos captured by the real portal.

The ability of a person to walk about the real building and for their presence in front of different portals to be reflected in Second Life addressed the vacancy problem to some extent by allowing the user to explore the real environment in a normal fashion but for their position to occasionally be updated in the virtual environment. However the abstract layout of the virtual environment presents a method of virtual environments exploration disjoint to the physical layout and relationship between the portals themselves, which is not ideal for all applications.

3.4.7 Doppelab

The current cross reality project undertaken by the Media Lab is the ongoing Doppelab, “*an immersive, cross-reality virtual environment that serves as an active repository of the multimodal sensor data produced by a building and its’ inhabitants*” [49, 50]. Doppelab uses the Unity3d game engine to allow users to visualise current and historic sensor data captured by a wide range of sensors distributed throughout the Media Lab in a three-dimensional virtual reconstruction.

However despite claiming to be a cross reality project, there is no evidence in the associated literature of any virtual-to-real, augmented virtuality data flow, which is a requirement of a cross reality system. Furthermore with regards to the vacancy problem, Doppelab “*focuses on encouraging and enriching individual users’ experiences of sensor data in 3-d environments*” and as such is of only passing interest to this review, as simultaneous presence in real and virtual environments must inherently support multiple users present in the same real and virtual surroundings to be of any great use.



Figure 3.11: The design of the virtual portals in the Ubiquitous Sensor Portals project.

3.4.8 Consumer Adoption of Cross Reality

A final piece of literature to consider whilst reviewing MIT's involvement with the cross reality paradigm and the vacancy problem, is an article written by Joshua Lifton, as an invited paper for the International Symposium on Ubiquitous Virtual Reality 2010, entitled *Consumer Adoption of Cross Reality*. This is a retrospective piece, written after Lifton had left MIT in 2007 to work at the Electric Sheep Company, “*then the premier third-party services company for designing, building and managing experiences across a number of virtual worlds, with a focus on Second Life*”, and in 2009 leaving the Electric Sheep Company to become an independent contractor and consultant. Lifton reasons that “*these experiences and the transition from academia to industry helped shape my views on what is possible and what is probable in the field of cross reality*”.

The overall conclusion of the paper is that cross reality hadn't yet achieved notable consumer adoption because the full utility of ubiquitous computing, particularly in terms of sensing and actuating infrastructure, and virtual worlds had yet to be realised as they hadn't been fully developed and deployed. As the key enabling technologies of cross reality, their full realisation is required before the paradigm can begin to be widely adopted.

As discussed in further detail in section 3.5.1, virtual worlds, particularly Second Life, failed to live up to their media hype, however the adoption and integration of sensor/actuator infrastructure into consumers' everyday lives has fared much better. Mobile phones with myriad sensors are now commonplace, on-body and off-body motion sensing have been extremely popular for games consoles (Nintendo Wii and Microsoft Kinect respectively), non-contact identification (RFID) is common in passports and building access and new buildings are now more often than not equipped with various sensing &/or actuating capabilities in the form of Building Automation Systems (BAS).

3.4.9 X-Reality

Marie Kim, Hwang Jae Gak and Cheol Sig Pyo at The Electronics and Telecommunications Research Institute, Korea, produced a Universal Sensor Network (USN) middleware, COSMOS, to address the issue of heterogeneity of sensor/actuator infrastructure that leads to difficulty in integrating them with virtual environments to establish a cross reality, or as they write it ‘X-reality’, system [19].

The matter of standards with relation to cross reality is discussed further in section 3.5.2 however Kim et al.’s article is mentioned here as its introduction section frames and explains the cross reality paradigm well in terms of the principle features

“The important point of X-reality is a paradigm shift from single-directional information flows to bidirectional information flows between two worlds.”

and also how it can be employed for simultaneous presence in real and virtual environments

“The differential characteristic of X-reality is that it can augment user’s engagement in the experiences of virtual presence and virtual world. Ultimately, it results in the human life span extension from the only real world to both worlds.”

Kim et al. also recognise the work at the time of Lifton et al. at the Media Lab for integrating sensing infrastructure with virtual worlds.

3.4.10 IEEE Pervasive Computing - Cross Reality Environments

The July-September 2009 issue of IEEE Pervasive (volume 8, number 3) was entitled ‘Cross Reality Environments’ and serves both as attestation to the research promise of the paradigm and as a good survey of the state of the art in cross reality research at the time of publication. The issue is introduced by Roy Want [12], whose definitions and matrix representation of the different categories of alternate realities were discussed in section 3.2.2, whilst the guest editor’s introduction is provided by Joseph Paradiso, director of the Responsive Environments group at the MIT Media Lab [18], whose work was discussed in section 3.4.2.

Paradiso’s introduction frames the background of cross reality very well, explaining the constituent technologies whose evolution will see the paradigm come to be, realising that it promises to serve as an extension of human perception and interaction and contrasting it with other alternate realities

“The genesis of ubiquitously networked sensor/actuator infrastructure, leveraged by the increasingly low cost of microelectronics, sensors, and wireless technologies, is endemic to pervasive computing. As the independent application-siloed sensor/actuator networks now being deployed begin to converge through common standards, the world is becoming covered by a seamless electronic “nervous system” that extends across places, things, and people. Concurrently, although 3D virtual environments have been around for decades, their deployment has recently exploded with the advent of massively shared online virtual worlds. Dedicated immersive games such as World of Warcraft and general purpose worlds like Second Life each now boast more than 10 million subscribers.

We call the ubiquitous mixed reality environment that comes from the fusion of these two technologies cross-reality. Sensor networks can tunnel dense real-world information into virtual worlds, where this data is interpreted and displayed to dispersed users. Interaction of virtual participants can incarnate into the physical world through a plenitude of diverse displays and actuators. We can envision a user’s interface into this environment as an extension of human perception and interaction, augmenting our five senses well beyond the canonical “here and now” and redefining the meaning of presence.”

Beth Coleman’s article [2] starts promisingly by further confirming that the terms cross reality and X-reality are interchangeable and then summarising the Eolus One project, an example application of cross

reality, as further discussed in section 3.4.11. However the other two example projects that Coleman discusses, the Green Phosphor 3D oil map and the Parsec voice-controlled virtual environment, lack the virtual-to-real, augmented reality information flow required for them to constitute cross reality systems under the criteria adopted by this review.

Lifton and other members of the Media Lab also have an article in the issue [51], which primarily serves to summarise their current cross reality endeavours (including Plug, Shadow Lab, Ubiquitous Sensor Portals, Tricorder and UbiCorder), but also provides an interesting background to and future visions of the concept from their perspective, leading to this somewhat colourful quote

“We see cross-reality precipitating when diverse and ubiquitous sensor and actuator networks meet pervasively shared online virtual worlds, where phenomena freely tunnel between real and contrived continua at a multitude of “wormholes” opened by densely deployed networked devices, seamlessly adapting the level of immersion to match a variable ecology of available interfaces and user context or preference.”

Dindar et al.’s article [52] discusses the importance of higher level inference techniques to make sense of phenomena detected by sensor networks, presenting it as an essential functionality of cross reality environments. This is a subject that Lifton touches upon in his thesis, where he discusses relationships such as vibration data inferring whether a lamp was plugged in before being turned on. Dindar et al present an example of such an event processing system, DejaVu, that works along with an example cross reality application that feeds real world RFID data into a virtual library in Second Life.

Horn et al. [53] discuss the scalability problems of the virtual worlds of the day (both ‘general purpose’ virtual worlds including Second Life as well as gaming worlds including World of Warcraft and Eve Online). They propose the Meru Project from Stanford as a virtual world architecture that can grow to global scales by modelling the virtual world after the real world; for example, using distance to determine the detail/bandwidth with which to render an object. Horn et al. also discuss the cross reality paradigm through the intriguing idea of ‘Space Zero’, a space within which physical objects can register themselves and thus advertise their presence and location to users of the virtual world. An example of such a relationship given is of a real person walking down a street in the real world, with an avatar’s position within a virtual world updated by GPS, allowing shop owners who have registered their shops with Space Zero to be able to advertise specials via the avatar. This discussion also leads to the valid and interesting point that, just like with the current Web, it won’t always be actual users observing the virtual world of a cross reality system, as it may well be scripts/bots that, for example, comb through advertisements of specials for those that match the user’s interests.

The issue of scaling virtual worlds to Internet scale is investigated in detail in Justin Clark-Casey’s masters dissertation [54]. Justin was then, and continues to be today, one of the lead developers of OpenSimulator.

The ‘Works in Progress’ section of the issue [55] discusses a number of projects under the title ‘Building and Employing Cross-Reality’ however according to the definition of cross reality adopted by this review only the first project, ‘6thSpace’, constitutes a cross reality system. 6thSpace was a 2009 project from Malden Labs that used Sun’s Project Wonderland virtual world (now the Open Wonderland project) to allow remote system administrators to collaborate with on-site staff to diagnose problems with computer equipment, with the remote participants gaining access to terminal windows on the on-site hardware by leveraging Wonderland’s application-sharing feature [55].

This same application-sharing feature has also been investigated by groups of scientists to remotely control laboratory equipment, visualising results in 3D by interfacing with external data sources, enabling collaborative lab-based research to be undertaken by geographically dispersed researchers [56].

3.4.11 Eolus One

A project from Swiss construction, building services and real estate company Implenia, involving creative minds from SAP, Wago and Zumtobel, as well as some crossover with IBM, that ran for 2 years exploring concepts including building automation, energy monitoring, alert management and preventive maintenance [2, 57]. Eolus One connected real-time data collection and distributed control mechanisms of smart

buildings in the real world to a Virtual Command Center (VCC) hosted in Second Life. This connection was facilitated by a hardware and software platform that they created and dubbed the Virtual World Communication Interface (VWCI), which mediated communication between Second Life and various protocols in Building Automation Systems (BAS) including ZigBee, CANOpen and Modbus.

The result was a 20 to 27 percent reduction in the energy consumption and carbon footprint of buildings under this control due to improvements in the manner in which data streams from multiple systems (heating, cooling, water usage, fuel consumption) could be combined, visualized, analysed, tracked and interacted with. The ability to visualise how multiple streams interacted with and affected each other allowed issues that were not apparent when studying the streams on an individual basis to be identified and resolved.

An additional benefit was a substantial reduction in the amount of manpower required to administer large numbers of buildings by allowing an individual operator to assimilate more information without the risk of missing important events during an ‘information overload’ than when using previous technologies.

3.4.12 IBM and Implenia

In 2008, during the height of their investigation into virtual worlds, IBM announced their 3-D Data Center technology to allow the recreation of data centres in secure virtual worlds, bringing real-time data from different facilities into a 3D environment to visualize hot spots, data flow, server utilization and more to reduce cost, save time and help reduce carbon footprints [58, 59].

Implenia made use of these solutions to extend the functionality of their existing VCC (see section 3.4.11) by using IBM’s virtual world integration middleware, Holographic Enterprise Interface (HEI), to add data from datacentre equipment to their existing virtual world models. Before the advent of this IBM technology Implenia’s knowledge of the state of their data centres comprised only of information from the BAS and their VWCI, with no information from the servers themselves.

“Until working with IBM we only knew the state of our data center from the information we got through the building automation system and our virtual worlds communication interface. We didn’t know the state of the server and information that was readily available to us until it was made more accessible via the 3-D visualizations that IBM built for us. We think that by combining this information with the information we had from the building automation side we can, from a building management standpoint, control the data much better and take action to be more efficient.” [59]

By introducing this information to their existing models, Implenia were allowed a finer control of the HVAC and security systems, allowing them to take action to be more efficient, with the consolidated views giving their operators insight into real physical issues such as how heat and energy flowed through the data centre.

3.5 Components of a Cross Reality System

As has been alluded to before in this review, a cross reality system relies upon the combination of sensor/actuator infrastructure with a virtual environment. As systems that support simultaneous presence in real and virtual environments will likely share many implementation details with cross reality systems, this section of the review provides a brief overview of the state of these constituent technologies.

3.5.1 Virtual Environments

Whilst inaugural cross reality projects and the majority that have since followed have made use of virtual worlds such as Second Life [60], OpenSimulator [61] and Open Wonderland [62], the paradigm can also be implemented using ‘traditional’ game and visualisation engines such as Unigine [63], Unreal Engine 3 [64] and Unity [65]; the previously mentioned Doppelab project from MIT (see section 3.4.7) has already made use of Unity for such a purpose.

The definition of ‘virtual world’ adopted by this review is that of ‘real virtual worlds’ presented by Yesha Sivan, in a 2008 article for the Journal of Virtual Worlds Research [6], as an aggregate of four factors; a three-dimensional world, community, creation and commerce. This definition is particularly useful as it differentiates between certain genres of games such as Massively Multiplayer Online Role-Playing Games (MMORPGs) including World of Warcraft (WoW) which share many features with virtual worlds; WoW presents players with a three-dimensional world, has a strong aspect of community by promoting players to join guilds to quest together and has a commerce aspect by allowing players to buy and sell in-game items. However WoW does not feature the creation aspect, as all content in the game is created by the game studio; in virtual worlds such as Second Life, the vast majority of the content is created by the users.

Several factors stand out in the literature for why, since their inception, virtual worlds have been used so extensively compared to game/visualization engines for alternate reality research. Foremost the accessibility of virtual worlds with regards to content creation makes them a more attractive option for researchers who may not have prior experience in a standalone 3D modelling package, which is a common requirement for working with game and visualisation engines. Virtual worlds provide powerful yet straightforward interfaces for building and personalising avatars, virtual spaces and objects and while these tools may lack some of the more advanced features and effects of professional modelling packages, they allow researchers to create diverse and expansive environments more easily and affordably, both in terms of experience and time required [66]. This allows more to be achieved within a limited budget as researchers can focus on running experiments and collecting and processing results rather than creating the experimental environment itself.

These experiments can be run for extended periods of time, over a period of months or more, at very low cost, thanks to the open licensing that many virtual worlds have [67]. Researchers are free to make use of existing hardware at their institution, deploying and administering their own infrastructure in any manner they see fit.

However virtual worlds are woefully outclassed by game engines when compared based upon technical merit rather than upon development and licensing costs. Modern game engines boast near photorealistic graphics, with incredibly complex lighting engines, physics engines capable of simulating soft and rigid body physics, cloth physics, rag doll physics, inverse kinematics and fluid buoyancy and interaction, sound engines supporting virtually unlimited 3D sound sources, Doppler effect, sound occlusion and multiple reverberation zones, artificial intelligence frameworks for non-player characters (NPCs) and environment editors capable of creating limitless, destructible environments with intelligent river and foliage generation and no loading when transitioning between locations.

Whilst all of these features can improve the user’s experience, at the cost of substantially longer, more complex and more expensive development time, it is the multi-platform compatibility that many modern game engines offer that represents their most lucrative feature over today’s virtual worlds. Many game engines allow the development of a single source that can be deployed onto desktop and laptop computers running Windows and Mac OS, games consoles and also to portable devices such as mobile phones and tablets running iOS and Android. This is a stark contrast to the current offering of virtual worlds, which require the use of a client program that only runs on desktop and laptop computers running Windows, Mac OS or Linux, drastically limiting the possibilities of virtual world usage on portable devices.

Virtual worlds, particularly Second Life, failed to live up to their media hype. In his retrospective on the consumer adoption of cross reality (see section 3.4.8) Lifton reasons that this was predominantly “*because it was too difficult to use and because it was never clear to most consumers what they should be doing once in Second Life*” [40]. The explosive popularity of social networking can also be attributed to the failure of virtual worlds to garner substantial followings, as they allow similar expression of identity and creation of social and business contacts in the more familiar and accessible environment of a Web browser rather than the comparatively alien and bespoke interface of a virtual world client.

The magnitude of this hype is wonderfully captured by an article published in Technology Review in 2007, in which the author Wade Roush, visionary of the Second Earth concept (see section 3.6.1) wrote “*according to technology research firm Gartner, current trends suggest that 80 percent of active Internet users and Fortune 500 companies will participate in Second Life or some competing virtual world by the end of 2011.*”

However this failure of virtual worlds to garner substantial consumer adoption does not represent the end of progress toward three-dimensional extension of the Web; indeed the fact that virtual worlds were developed at all and for a while gained vast media attention and hype, and that complex 3D applications can now be rendered within a Web browser via extensions such as the Unity Web Player, is proof that the concept of the 3D Web stands to succeed. Once infrastructure is put in place that allows rich 3D content to be accessed as easily as the 2D Web can currently be accessed via any Web browser, concepts such as cross reality and simultaneous presence in real and virtual environments can emerge and thrive. Even though game engines that allow compilation of binaries for multiple platforms represent greater accessibility than today's virtual worlds, the installation of a separate program or plugin for the sole purpose of accessing a single 3D environment is still a substantial inconvenience when compared to the ability of a Web browser to access any Web content.

3.5.2 Sensor/Actuator Infrastructure

Numerous platforms for distributed sensing and actuating exist [8]. Commercially procurable hardware designed specifically for the task, such as wireless sensor network (WSN) solutions including Berkeley motes [8], are one option whilst general-purpose low-power prototyping and embedded platforms, such as Phidgets and Arduino, have also proven popular for creation of bespoke sensor/actuator devices [68]. Software platforms include those tailored specifically for operation on WSN motes such as TinyOS [69] and more general purpose operating systems such as Contiki [70] which is designed for running on devices with limited memory footprints and computational capabilities. With sufficient expertise and resources bespoke platforms can be constructed from scratch, as the Media Lab did for the creation of their Plug platform, by selecting an appropriate microcontroller and designing and manufacturing a node around it, as well as writing the necessary operating system and software to run atop the resulting bespoke platform.

Many authors, including Lifton in his inaugural thesis, have recognised that the emergence of cross reality will depend upon sensor/actuator infrastructure as ubiquitous as today's electrical sockets and lighting fixtures. However whether this provision will come from the wider installation of dedicated sensor/actuator infrastructure into our environment, for example through continuing adoption of Building Automation Systems and 'smart' building construction, or from integration with sensor/actuator infrastructure that is now commonly being integrated into consumer electronics devices such as mobile phones, tablets and games consoles, or indeed from a combination of both, is yet to become apparent.

But as Paradiso observes in his guest introduction to IEEE Pervasive 8.3 (see section 3.4.10) the ability to access such sensor/actuator infrastructure via common standards will be one of the enabling features for widespread deployment and adoption of cross reality and systems that support simultaneous presence in real and virtual environments. This likely will not come via the global adoption of any single hardware &/or software platform, but by the heterogeneous array of different hardware &/or software platforms adopting a single standard for abstracting over their access and control interfaces.

A number of standards and frameworks that attempt to provide such a standard method of interfacing with sensor/actuator infrastructure have already been developed, several of which will be discussed in the following part of this section of the review. It is prudent to appreciate that each was designed and developed from a different perspective and with different final goals. Whilst none are explicitly designed for facilitating implementation of alternate reality projects, they represent substantial work toward a standard interface to heterogeneous sensor/actuator infrastructure, which would facilitate the easier creation of cross reality systems and thus systems that support simultaneous presence in real and virtual environments.

Open Geospatial Consortium Sensor Web Enablement

The Open Geospatial Consortium (OGC) is an international standards organisation comprising more than 400 industry, academic and government institutions collaborating in the development and implementation of open standards for geospatial and location services. The Sensor Web Enablement (SWE) project [71] is a collection of such standards for exploiting Web-connected sensor systems.

SWE aims to enable all Web and/or Internet accessible sensors (including imaging devices such as surveillance cameras) to be accessible and controllable via the Web, through open Web standards, creating in effect a ‘sensor Web’ that can be accessed in a similar manner as to how HTML and HTTP allow access to the WWW. As of 2008 SWE was in use by projects from Oak Ridge National Laboratory, the Defence Intelligence Agency, the European Space Agency and NASA, amongst others.

SWE is comprised of 7 individual standards, many adopted as OpenGIS standards, each concerned with a different aspect of Web-based interaction with sensor networks. Together they facilitate the discovery (with no *a priori* knowledge) of sensor systems, including their properties, geospatial location and capabilities, the requesting, filtering and retrieving of observations from these systems, the encoding of these observations for communication across a network and the publishing of and subscription to alerts issued by sensors when certain conditions are met (such as particular environmental conditions).

SWE also harmonizes with other relevant sensor and alerting standards including IEEE 1451 (discussed later in this section of the review), the OASIS Common Alerting Protocol (an XML based data format for exchanging public warnings and emergencies between alerting technologies), the OASIS Asynchronous Service Access Protocol (standard descriptions of how to start, manage and monitor long running services) and Web Services Notification (specifications related to the WS-Resource framework that allow event driven programming between web services).

SWE was demonstrated in December 2006 at the fourth OGC Web Services testbed activity, OWS-4, using a hypothetical scenario of a bomb containing radioactive material detonating near New York City (this demonstration and others are available to view at [72]). SWE was used to find, access and integrate diverse geospatial resources from the vicinity of the detonation, many of which were live sensors, with all of the information flowing from different sources through Web services. First a radiation sensor triggered an alert which, via subscriptions, caused a chain of services to operate; other sensors in the vicinity were polled, surveillance cameras near the area became available to watch and position and NASA Earth Observation satellites were instructed to provide images of the area from orbit over the next few days so as to document the movement of the radioactive plume.

This was a demonstration of a Geospatial Decision Support System (GeoDSS) which allows a decision maker at a single workstation to identify and access geospatial resources no matter where they are, to bring them into an operational context and integrate them with other resources to support the overall decision making process.

As its name implies, SWE is primarily focussed on the provision of interfaces that exploit open Web standards to access and administer geospatially distributed *sensors*. Whilst the standards provide methods for transmitting control information to these sensors, these are concerned with administration of sensors and are not intended to be used for controlling complex actuating devices. Whilst SWE can provide a cross reality developer with extensive infrastructure for implementing the augmented virtuality part of a cross reality system, it would need extending to provide sufficient infrastructure to allow the augmented reality part to be implemented successfully to the same degree.

Ubiquitous Sensor Networks

A Ubiquitous Sensor Network (USN) is a conceptual network built over existing physical networks which makes use of sensed data and provides knowledge services to anyone, anywhere and at any time, and where information is generated by using context awareness. USNs can utilize both traditional wired and wireless networks as well as Wireless Sensor Networks (WSN) [73].

Whilst initial WSN development focused on isolated networks that captured and transmitted sensed data to designated application systems, the continued development of the paradigm has led to the current ability to build an intelligent infrastructure of sensor networks connected to *existing* physical networks. This information infrastructure has been dubbed Ubiquitous Sensor Networks, due to the introduction of sensors to pre-existing ubiquitous network infrastructure, and opens possibilities for applications and services based on sensor networks to various customers, from consumers and public organisations to enterprises and governments. Deploying sensor/actuator infrastructure upon existing network infrastructure promises to

reduce installation and operational costs, as well as allowing to a certain degree the leverage of existing network standards for the access and control of the sensor/actuator infrastructure.

USN applications are therefore created by integrating sensor network applications and services into existing network infrastructure. This integration may in some cases require extensions and/or additions to core network architectures in order to cover the functional capability requirements extracted from USN applications and services. Thus the requirement for a USN middleware was recognised, to provide the necessary interface sets and functions for USN applications and for sensor networks to be easily integrated [19].

The importance of this middleware is in the decoupling of USN application implementation and sensor network implementation, such that any physical sensor hardware from any vendor can be integrated with an existing network to produce an USN. This approach makes sensor networks a viable and cost effective endeavour, when manufacturers are able to sell the same sensor networks to different service providers to achieve different aims, rather than having to build bespoke platforms for different applications.

The importance of supporting USN applications and services in the Next Generation Network (NGN) environment has been recognised and addressed by the International Telecommunications Union (ITU) with the publishing of Rec.ITU-T Y.2221 [73]. These recommendations cover administration and management of sensor networks, profile management, registration and discovery, quality of service (QoS) considerations, support for location-based service support, mobility support, security, and so forth.

Kim et al. recognised such a USN middleware as the key technology to propel the realization of the cross reality paradigm, by providing a solution to the inherent heterogeneity of sensor and RFID devices and the data that they produce. In their 2009 paper [19] discussed in section 3.4.9 they present an implementation of an USN middleware COSMOS that intermediated between various USN applications and simulated physical resources such as RFID reader networks and sensor and actuator networks, then showing how such a middleware functions as part of a larger cross reality platform.

Another USN middleware from Korea is XCREAM (XLogic Collaborative RFID/USN-Enabled Adaptive Midddleware) from a collaboration between Towson University, USA, and Sangmyung University, Korea, which aims to allow collaboration between many RFID/USN applications by providing them with a flexible interface through a web-based service scheme, XML-based infrastructure and the XLogic script language [74].

Microsoft SenseWeb

SenseWeb from Microsoft Research is a system that allows peer production of sensing applications, producing new kinds of media and applications over existing data networks, by allowing users to grant access to their sensors to other remote users. The basic premise is that contributors deploy their own sensors or sensor networks, which might be designed for their own dedicated application or simply for sharing with other users, uploading the observations from these sensors to the SenseWeb system where they can then be accessed by other SenseWeb users through an application-specific GUI. Using SenseWeb, applications can initiate and access sensor data streams from shared sensors across the entire Internet [75].

The approach of the system is such that applications can access resources contributed by any entity in a uniform manner, regardless of the underlying platform of that entity. To achieve this, sensors are connected to a sensor gateway which provides a uniform interface to all components above it, hiding much of the complexity of the heterogeneous communication protocols, power capabilities, bandwidth capabilities, etc. of the numerous sensor platforms beneath.

However unlike with systems that implement IEEE 1451 compliant smart transducers (see section 3.5.2), when a developer wishes to connect sensors to a sensor gateway using SenseWeb she must manually define many of the characteristics and properties of these sensors; type, location (longitude, latitude, altitude), description, etc. Additionally for the observations from different sensors to be made useful (eg two different types of temperature sensor, one that reports observations simply as a voltage from a thermistor and the second that performs some on-sensor processing and calibration and reports observations as a value in centigrade) a ‘data transformer’ must be used to convert data semantics through processing. Whilst this approach to using intermediary data transformation allows for powerful handling of data, such as passing a video observation through a people count algorithm and only presenting to other users the output of this algorithm rather than the video itself, it adds complexity and manual implementation to accessing

and making sense of observations which could have been achieved via a system such as the TEDS of IEEE 1451. When a user wishes to connect sensors to a sensor gateway, there is no automatic discovery by the sensor gateway of these sensors, nor of their characteristics (available sensors/actuators, calibration, type of communications interfaces, etc.) [76].

Communication between sensors and sensor gateways relies upon the existence of drivers. Whilst many common sensing devies, including wireless motes and network cameras, already exist, this approach means that scalability to new and particularly unusual or unpopular (in terms of widespread adoption) sensing platforms is poor. Whilst SenseWeb provides some very powerful techniques at the application interface level for abstracting over heterogeneity at lower levels, such as differences in requirements to conserve power and bandwidth, by making extensive use of caching and distribution of tasking to different individual sensors when many are available that can achieve the same observation, all of this abstraction assumes that the sensors themselves have already been integrated with the system, usually requiring manual human intervention.

SenseWeb presents a very polished interface to the end user, making it extremely easy and intuitive to pull in observations from sensory resources around the globe and put them to good use, however this ease of interaction at the end user level comes at a price lower down where sensor attributes must sometimes be manually specified and intermediary processing must sometimes be implemented. SenseWeb serves as a good example of what the interface to a standardised framework for cross reality development should be like, however automatic resource discovery and administration at the lower levels would make the system more desirable to developers at all levels, rather than just at the receiving end of observations.

Global Sensor Network (GSN)

GSN is a middleware developed by École Polytechnique Fédérale de Lausanne (EPFL) to support the rapid and simple deployment of a wide range of sensor network topologies, facilitate the flexible integration and discovery of sensor networks and sensor data, enable fast deployment and addition of new platforms, provide distributed querying, filtering and combination of sensor data and support dynamic adaptation of the system configuration during operation [77].

EPFL researchers realised that as the price of wireless sensors decreased and the number of autonomous sensor networks increased, the emerging desire for a ‘Sensor Internet’ was being hampered by the difficulty of sharing and integrating data among heterogeneous sensor networks, as previous research in the sensor network domain had mainly focused on routing, data aggregation and energy conservation inside a single sensor network.

After identifying the standard strategy adopted successfully in other domains of addressing such a problem by devising a middleware which provides powerful abstractions allowing easy integration of concrete physical platforms, EPFL set about developing GSN.

GSN’s key abstraction are ‘virtual sensors’, which abstract from implementation details of access to sensor data and correspond either to a data stream received directly from a sensor or to a data stream derived from other virtual sensors. Virtual sensors can be any kind of data producer; actual sensors, wireless cameras, desktop computers, mobile phones, or even a combination of these. Each virtual sensor can have any number of input streams but produces exactly one output stream based on these input streams and some logical processing.

Virtual sensors are described in a declarative deployment descriptor, which is an XML file defining various attributes; a unique identifier, any number of input streams, the structure of the output stream, etc. Defining this XML file is all that is required to deploy a new sensor as long as GSN includes software support for the concerned hardware/software. This support is provided by wrappers conforming to the GSN API and several such wrappers are provided to cover common hardware/software including a generic HTTP wrapper for accessing data from devices using HTTP GET and POST, and a serial forwarder wrapper that supports interaction with TinyOS WSN motes over a direct serial connection. Implementing new wrappers is not a substantial endeavour either, as even complex wrappers are only a few hundred lines of code (mostly Java).

GSN has been demonstrated on real hardware through experiments with typical high-load application profiles and is used as the primary streaming data database for the ETH Centre for Competence Environment and Sustainability (CCES) and through the Swiss Experiment platform project the usage of GSN is

increasing, currently in the region of 80 – 100 million data points. Furthermore, GSN is open source under the GNU GPL and available freely on the Web, making it ideal as a platform for extending upon to better cater to specific project needs.

IEEE 1451 Networked Smart Transducers

IEEE 1451 is a family of Smart Transducer Interface Standards defining a set of open, common, network-independent communication interfaces for connecting transducers to microprocessors, instrumentation systems and control/field networks [78]. The main goal of IEEE 1451 is to allow network access to standardised transducer data through a common set of interfaces, whether the transducers are connected to systems or networks via wired or wireless means [79].

A ‘transducer’ is a device that converts energy from one form to another and may be thought of as a super set of sensors and actuators; a sensor is a transducer that generates an electrical signal proportional to a physical, biological or chemical parameter (such as an increasing voltage in response to an increasing level of ambient light), whilst an actuator is a transducer that accepts an electrical signal and takes some physical action (such as switching an electrical relay in response to a certain electrical signal). A ‘smart’ transducer extends the concept by integrating an analog or digital sensor or actuator element (a transducer) with a processing unit and a communication interface [80]. A smart transducer implementation thus comprises a hardware or software device consisting of a small, compact unit containing a sensor or actuator element, a microcontroller, a communication controller and associated software from signal conditioning and calibration to diagnostics and communications [81].

An IEEE 1451 compliant smart transducer specifies a Transducer Electronic Data Sheet (TEDS) and partitions the system into two major components; a Network Capable Application Processor (NCAP) and a Transducer Interface Module (TIM), the two connected by a Transducer Independent Interface (TII). The NCAP is a network node which performs application processing and network communication functions, while the TIM consists of transducer signal conditioning and data conversion and a number of sensors and actuators themselves (up to 255, in any combination). In this fashion many sensors and actuators can be hosted by a single TIM and many TIMs may communicate with a single NCAP via the TII which defines a communication medium and a protocol for transferring sensor information.

From a standards perspective this approach helps to achieve sensor ‘plug and play’ and interoperability for industry and government [79] whilst the distribution of measurement and control can dramatically reduce network load by allowing transducers themselves to determine whether observations are worthwhile communicating and also to encode them into formats more suited to the network being employed before transmitting [82].

This standardisation of communication with and between transducers will be extremely helpful for a standardised support framework for cross reality development. Instead of such a framework having to support numerous different transducer interfaces to cater for the numerous different hardware and software platforms on the market, or even different interfaces for sensors as for actuators, all communication between framework and transducers can be achieved via IEEE 1451, allowing the framework to support any hardware/software platform that adheres to the standard.

OGC SWE (discussed in section 3.5.2) adopts a similar approach by supporting interaction with IEEE 1451 compliant smart transducers through the standard interfaces presented by their implementations of the NCAP. This allows SWE to provide access to these transducers via interfaces implemented through open Web standards.

3.5.3 Integration of Sensor/Actuator Infrastructure with Virtual Environments

With the exception of IEEE 1451, none of the standards, frameworks and middlewares discussed in the previous section have explicit support for controlling actuators; that is, they are primarily concerned with administering and retrieving data *from* sensors with little or no attention or support provided for sending commands *to* actuators. A cross reality system and any system in which there is simultaneous presence within real and virtual worlds will naturally require this bidirectional flow of control and media.

Additionally the standards, frameworks and middlewares covered in the previous section are all from the general perspective of sensors &/or actuators and are not explicitly to do with alternate realities, cross reality or simultaneous presence in real and virtual environments. Whilst the argument holds that a well designed and widely adopted set of standards-compliant interfaces to ubiquitous sensor/actuator infrastructure is all that is required to be able to combine such infrastructure with virtual environments and create cross reality systems capable of supporting simultaneous presence in real and virtual environments, there also exist standards that cater specifically for the mediation of sensor data and actuator control to and from virtual environments, as well as between different environments, by defining encoding formats for sensor data and actuator commands, communication standards, etc.

MPEG-V

The most comprehensive work to date toward a set of standards for interaction between the real world and virtual worlds, and between different virtual worlds, is MPEG-V, a work that started in 2007 as a unified effort to develop standards within virtual worlds and between virtual worlds and the real world. The culmination of this work was the publication of ‘ISO/IEC 23005 Information technology - Media context and control’ in January 2011. To quote directly from the first section of the standard

“ISO/IEC 23005 (MPEG-V) provides an architecture and specifies associated information representations to enable the interoperability between virtual worlds, for example, digital content provider of a virtual world, (serious) gaming, simulation, DVD, and with the real world, for example, sensors, actuators, vision and rendering, robotics (e.g. for revalidation), (support for) independent living, social and welfare systems, banking, insurance, travel, real estate, rights management and many others.” [83]

Yesha Sivan, head of the information systems programme at the Tel Aviv Academic College, brought the idea of standardisation in the field of virtual worlds to the ITEA2 project outline event in Dusseldorf in 2007, gaining the attention of Jean Gelissen from Philips Research who teamed up with Sivan and took on the role of project leader for Metaverse1, one of the largest contributors toward MPEG-V.

Sivan’s call for participation in MPEG-V was published in the November 2008 issue of the Journal of Virtual Worlds Research (JVWR, of which he now acts as Managing Editor) and appropriately titled ‘Real Virtual Worlds SOS (State of Standards)’ [84]. Here he makes the observation that the virtual worlds industry of the time was operating more akin to the computer gaming industry than the Internet industry, in that each virtual world had its own client and server that were incompatible with other virtual worlds whereas a Web browser can navigate to any Web page. At the time of this article’s publication, MPEG-V was a working group under the ISO (ISO/IEC JTC 1/SC 29/WG 11/N9902).

The October 2009 issue of JVWR was dedicated to ‘Technology, Economy & Standards’ and devoted several articles to the progress toward release of MPEG-V, starting with an editorial from Sivan giving an overview of the state of virtual worlds standards in 2009 [85]. Linden Lab themselves (the creators of Second Life) contributed a short article to this release of the journal [86], expressing their opinion that virtual world companies would be most successful by providing only the minimal scaffolding for the development of the rich content experiences that would bring more and more usage and to confirm that where feasible their path had been to open the avenues for others to develop and explore the interconnected components that comprise their virtual world.

Gelissen provided an overview of MPEG-V as it stood in late 2009, explaining the role of the Metaverse1 project and of particular interest to this review noted that “*a strong connection between the virtual and the real world is needed to reach simultaneous reactions in both worlds to changes in the environment and human behaviour*” [87] which supports the notion that MPEG-V may be of use for developing systems for simultaneous presence in real and virtual worlds and addressing the vacancy problem.

After the release of ISO/IEC 23005 in its first edition, JVWR released an issue dedicated to ‘MPEG-V & Other Virtual Worlds Standards’. As Sivan notes in his editorial to this issue [88], it takes an historical perspective as in many ways it is a direct descendent of the October 2009 issue, which has just been discussed.

Gelissen and Sivan co-authored the Issue Editors' Corner introducing the special issue [89] and then went on to write a historic account of Metaverse1 and its input to the MPEG-V project [90]. The project involved about 30 EU based organisations, totalling around 100 people, and ran for 3 years from mid 2008 to mid 2011, resulting in around 216 person-years of work that was piped into the MPEG-V endeavour. They also update the assumptions that they made in 2008 and in particular note that the full impact of virtual worlds has not come to be, at least partly due to the rise of social media and mobile media grabbing the attention, an opinion shared by other authors. A use case of the Metaverse1 work was presented by Cabello et al. [91], in the same issue.

An overview of the areas in which MPEG-V applies in relation to virtual worlds and the real world is included as figure 3.12.

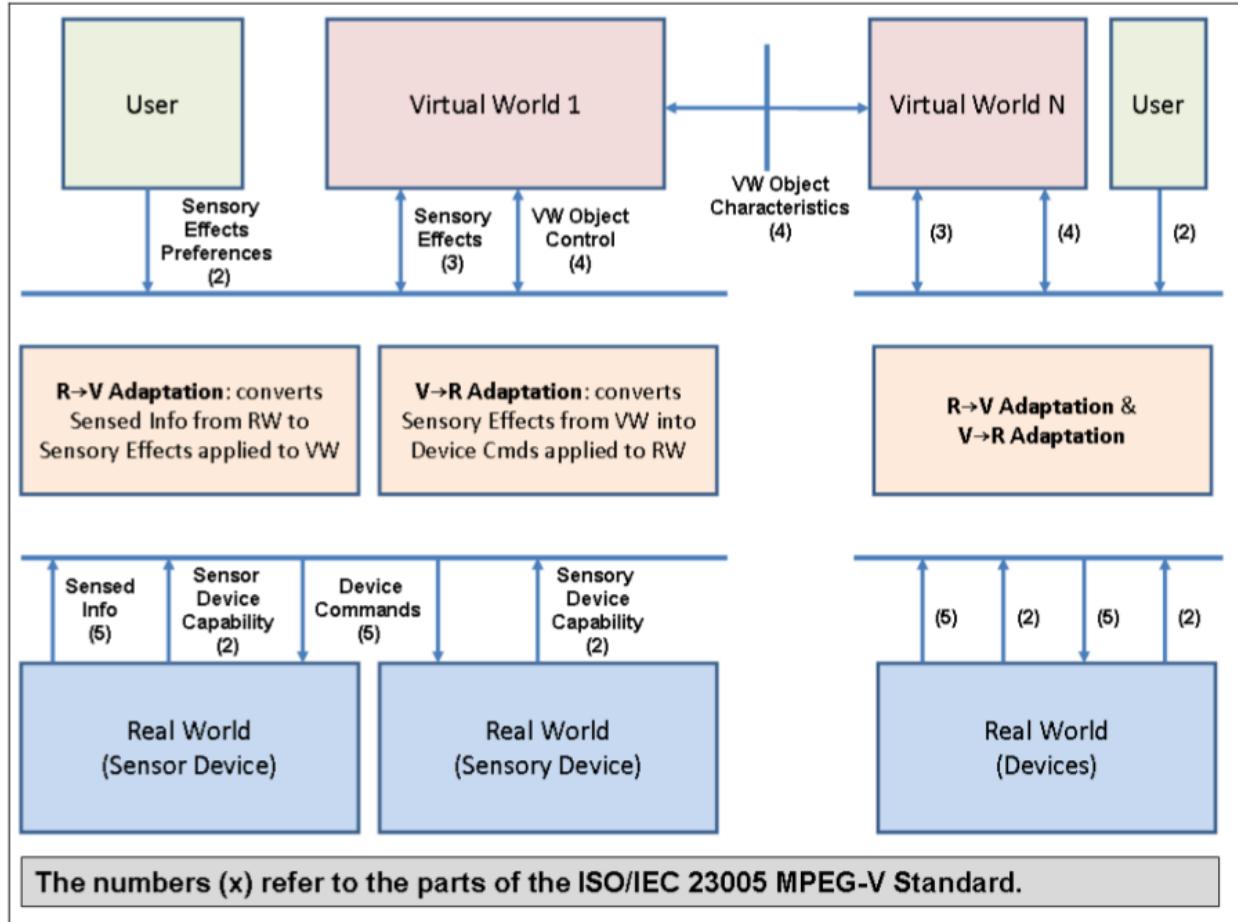


Figure 3.12: MPEG-V overview as of December 2011.

IEEE P1828 - Standard for Systems with Virtual Components

Due to possible licensing issues associated with MPEG-V, with some people envisaging a similar situation as with MP3 where in September 1998 Fraunhofer IIS sent a ‘letter of infringement’ to several small commercial and open source MP3 development projects, claiming that due to patents related to MP3 and held by Fraunhofer, all encoders (even those given away free) required the payment of royalties, a truly ‘open’, patent-free and zero litigation alternative to MPEG-V, the IEEE Standard for Systems with Virtual Components,

is under development as IEEE working group P1828 [92]. This endeavour is still in the concept phase, with a draft set for completion before the end of 2012. From the current draft document

“This standard is an over-arching standard for virtual environments. The over-arching standard provides a reference model and covers subsequent standards for virtual worlds. This standard establishes terminology for the Virtual Environment (VE) components and systems. The reference model covers addresses, interfaces, and/or communication protocols between and among system components. Components include supporting devices, like servers, reality engines/environments, virtual objects including avatars, and scripts. This standard defines processes, conventions and protocols to promote interoperability objects between and among virtual environments and networked resources. This standard is application agnostic but will provide illustrative use cases.”

“The purpose of this standard is to promote leverage (software and data reuse) of work between virtual project developments and promote shared resources. Another purpose of this standard is to establish the means to incorporate the respective strengths of different virtual environments. The purpose is to define a comprehensive abstraction of functions that virtual worlds can implement to interoperate at negotiable levels in near-real-time. [93]

One of the use cases already presented in the draft, under the title ‘Augmented and Mixed Reality’

“...covers a virtual environment with interfaces with real devices. Real devices will report status to include geolocation that would be reflected in world. Real devices will also respond to control from in-world avatars and/or scripts.”

Which clearly alludes to both virtual-to-real, augmented reality control and real-to-virtual, augmented virtuality control, supporting the standard’s suitability for enabling cross reality systems and/or systems that permit simultaneous presence in real and virtual environments.

Ultimately the realisation of cross reality and support for simultaneous presence in real and virtual environments will likely come through the development and deployment of both of these concepts; combining both an easy way to access and control sensor/actuator infrastructure via a standardized abstraction and an easy way to integrate this access and control into a virtual environment through standardized interfaces and encodings.

3.6 Use Cases

This section presents two use cases for simultaneous presence in real and virtual environments and illustrates the two different relationships that can exist between the real and virtual environments - whether they are *spatially equivalent* or not.

3.6.1 Virtual Time Windows - Spatial Equivalence

The ongoing Virtual Time Window (VTW) project is an application of simultaneous presence in real and virtual environments within the domain of cultural heritage that promises to further existing alternate reality work in the field by allowing simultaneous exploration of a real cultural heritage site and its virtually reconstructed counterpart via a tablet computer [94]. A visitor to a cultural heritage site, such as the ruins of the cathedral at St Andrews, uses a tablet computer that in effect presents a ‘window’ into a virtual reconstruction of the entire site as it was at an earlier point in time, such as the cathedral in its 13th century splendour.

To maintain a natural and unhindered sense of exploration VTW does not require visitors to manually control navigation within the virtual environment. Changes in the tablet’s position within the site are

automatically reflected by a corresponding movement of the avatar within the virtual environment, making use of a combination of location tracking technologies. The direction that a visitor faces is monitored by magnetometer and the angle that they hold the tablet at by accelerometer; this information is reflected by the direction and pitch of the camera within the virtual environment. The resulting style of interaction is similar to using a digital camera to take a photo; the screen on the back of the camera shows what the image will look like when the shutter is released, whilst with VTw the screen on the tablet shows what the site looked like in the past.

This approach addresses the vacancy problem by presenting the user with a convenient and natural manner in which to interact with the virtual environment whilst simultaneously exploring the real environment. A tablet computer is small, light and easy to carry and by controlling the position and direction of the camera by sensing the user's physical movements the user doesn't have to pay close attention to manually navigating within the virtual environment which would risk introducing vacancy from the real environment.

By using a complete virtual environment, rather than adding sparse virtual augmentations to a user's view of the real location in an augmented reality fashion, interactions between users at the site and those who are physically elsewhere are possible via the virtual environment.

With VTw, the virtual environment is based upon the real environment. Even though certain features may differ, for example where ruins stand in the real environment a complete building may stand in the virtual environment, there is a fundamental spatial equivalence between the two environments - they are both in effect the same 'location' or 'place', in an abstract sense of the terms. It is this relationship that permits the project to map a user's physical position in the real environment to an equivalent position within the virtual environment, allowing them to navigate both when in effect only controlling their navigation in one.

One might consider the 'Second Earth' concept to be the ultimate realisation of this scenario of spatially equivalent real and virtual environments. Discussed by Wade Roush in a 2007 article of MIT's Technology Review magazine [95], which is cited by Lifton in his thesis, Second Earth is theorised as the combination of the notions of virtual world technology (as in Second Life) with 'mirror world' technology (as in Google Earth); Second Earth theorises a virtual simulation/reconstruction of the entire physical world, such that for any location in the real world there is a corresponding location in the virtual world. Such a resource would allow for simultaneous presence in corresponding real and virtual environments to take place anywhere, rather than being restricted to specific real world locations for which a corresponding virtual location had been created, such as a cultural heritage site.

Furthermore, if one were to apply the concepts of cross reality to such a global virtual reconstruction, it would in effect create a complete parallel virtual Earth that would react in real-time to events in the real world via sensor infrastructure and would be able to affect the real world in real-time through actuator infrastructure.

Naturally the Second Earth concept will remain just that - a concept - likely for decades, as the underlying technologies and infrastructures are not yet available to us. In a blog post on the subject of virtual world/mirror world mashups [96], Avi Bar-Zeev estimates that the Second Life server model at the time would require 2.4 billion physical servers to host a simulation of the entire surface of the Earth, or 1400 servers just for Manhattan. In a comment on this post, Roush emphasises that his article in Technology Review was not meant to be taken as a premise for a literal Second Life/Google Earth mashup, but that they were the leading virtual world/mirror world technologies at the time and that overlap was sure to happen.

3.6.2 Snow Crash - No Spatial Equivalence

In the opening quote to this review, taken from Neal Stephenson's cyberpunk novel *Snow Crash*, the protagonist enquires about the location of another character, called Y.T., both in the real world and in the 'Metaverse'. For the sake of this discussion, this Metaverse can be considered analogous to a virtual world akin to Second Life, accessed via a head mounted display, and comprises an entirely synthetic virtual world whose locations have no counterparts in the real world. Y.T.'s response is that "*In the Metaverse, I'm on a plusbound monorail train. Just passed by Port 35.*" whilst in reality she is at a "*Public terminal across the street from a Reverend Wayne's*".

In this scenario there is no spatial equivalence between the real environment and the virtual environment - they are not the same 'location' or 'place' as is the case with VTW - however the concept of simultaneous presence in both can still be useful, as illustrated later in the book. Y.T. is in fact waiting for a third character, Ng, to come and collect her, which leads to the following conversation in the Metaverse between Y.T. and Ng

"... you're driving?"

"Yes. I'm coming to pick you up - remember?"

"Do you mind?"

"No," he sighs, as if he really does.

Y.T. gets up and walks around behind his desk to look.

Each of the little TV monitors is showing a different view out his van; windshield, left window, right window, rearview. Another one has an electronic map showing his position: inbound on the San Bernardino, not far away.

"The van is under voice command," he explains. "I removed the steering-wheel-and-pedal interface because I found verbal commands more convenient. This is why I will sometimes make unfamiliar sounds with my voice - I am controlling the vehicle's systems."

Ng is driving his van in the real world to come and collect the real Y.T., whilst simultaneously sitting in his virtual house in the Metaverse having a conversation with the virtual Y.T., using a series of virtual TV monitors in the Metaverse to inform him of his real surroundings and to allow him to control the van.

In this scenario there is necessarily no spatial equivalence between the real environment and the virtual environment, as the virtual environment has no spatial equivalent in the real world - the monorail and Ng's house have no real counterparts. However a lack of spatial equivalence between the real environment and the virtual environment could also arise when the virtual location does have a real counterpart, but the user isn't there. For example, in reality a user could be in London whilst in the virtual environment they could be in a location that is spatially equivalent to a real part of Hong Kong.

It is already common to see people interacting with their real location, even if that interaction is limited to just walking from A to B without bumping into too many other people or being run over by a bus, whilst simultaneously interacting with the 2D Web, social media and textual chat via mobile devices such as mobile phones and tablets. With the continued trend toward the 3D Web, it is no leap to imagine a near future in which people regularly want to interact with a 3D virtual environment that is not spatially equivalent to their current location at the same time as walking to the bus stop, thus simultaneous presence in real and virtual environments that are not spatially equivalent promises to be a desirable scenario.

3.7 Conclusions

We are rapidly approaching a situation in which ubiquitous sensor/actuator infrastructure allows us to access vast amounts of information about any location at any time and additionally to act upon this information and affect these locations. The continuing adoption of fast Internet connections, the increasing ability of commodity hardware including portable devices such as mobile phones and tablets to render complex three-dimensional graphics, and the development of 3D multi-user virtual environments that place an emphasis on notions of community, creation and commerce instead of competitive gaming, all point toward a continuing natural progression toward 3D extension of the Web on a large scale.

It is already common to see people spending substantial amounts of time immersed in the 2D textual/graphical Web whilst simultaneously interacting with the real world around them. This desire to

maintain a Web presence whilst simultaneously interacting with the real world is set to remain as interaction with the Web evolves from 2D to 3D. Thus it is prudent to investigate approaches for implementing and applications for exploiting the concept of simultaneous presence in real and virtual 3D environments, whether spatially equivalent or not.

This review has unearthed a plenitude of research on numerous alternate realities, either experienced in isolation from other realities (reality, virtual reality) or by mixing limited amounts of one with another (augmented reality, augmented virtuality) however has discovered a comparative lack of research attention focussed on the concept of simultaneous presence and interaction with two complete environments, one real and the other virtual. Cross reality is the closest existing concept, however the vast majority of the research in this field has used statically located computers to access the virtual environments, preventing users from exploring or interacting with the real environment that is not immediately surrounding them; the true notion of simultaneous presence in real and virtual environments requires freedom of movement and interaction with both environments, perhaps by adopting a manner of interaction with the virtual environment similar to that of the VTW project.

This deficiency of research into simultaneous presence in real and virtual environments warrants addressing with further academic investigation, as it represents a style of interaction that is bound to become commonplace as progress toward 3D extension of the Web continues at an accelerated pace.

4

Virtual Time Window

5

Mirrorshades - Design/Implementation

5.1 Introduction

This research centres 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*” [14].

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 [14, 49]. 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’ [97].

5.1.1 Position - *Milgram & Kishino*

The position of XR in relation to other alternate realities studied by Computer Science can be visualised using Milgram & Kishino’s *virtuality continuum* (figure 5.1) [10] that stretches from an entirely real environment at one extreme to an ontologically parallel but entirely virtual environment [98] at the other. The explanation herein distinguishes between environments themselves (depicted in figures 5.2 to 5.5 by solid ellipses) & where the stimuli that the user is perceiving originate from (depicted by dashed ellipses).

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 (figure 5.2), a XR system instead features two discrete environments, one real & the other virtual, each complete unto itself (figure 5.3).

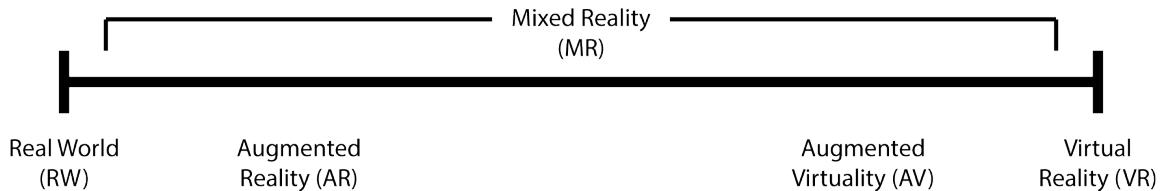


Figure 5.1: Milgram & Kishino's virtuality continuum.

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 (figures 5.4 & 5.5) in a manner that allows them to engage with each environment without becoming wholly vacant from the other.

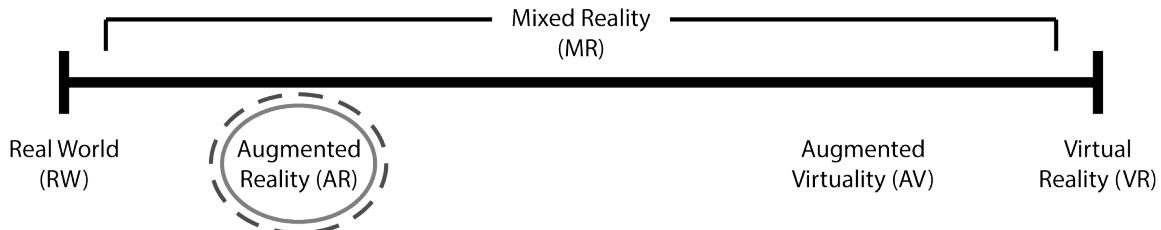


Figure 5.2: AR visualised using the virtuality continuum.

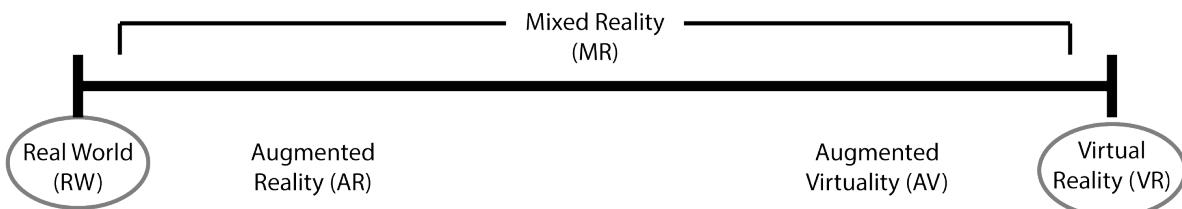


Figure 5.3: The two environments that comprise a XR system.

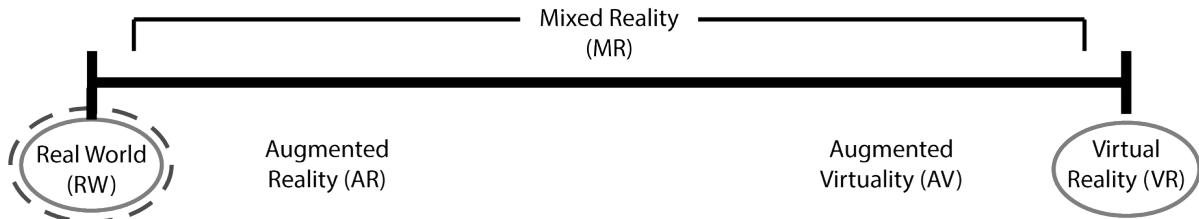


Figure 5.4: A XR system with the user attending to RW stimuli.

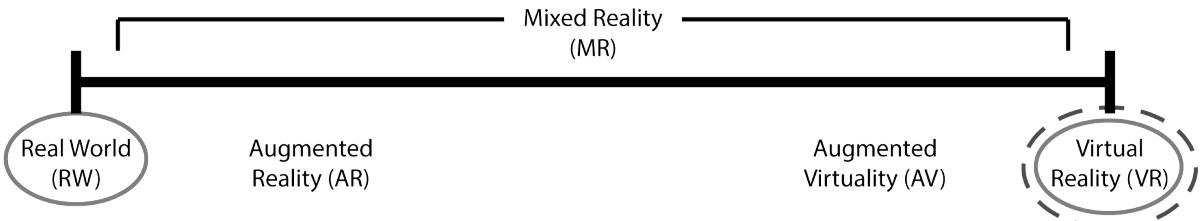


Figure 5.5: A XR system with the user attending to VR stimuli.

5.1.2 Position - Waterworth & Waterworth

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 [99]; the combination of these models is shown by figure 5.6. 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 presence¹ in that environment toward its antithesis – absence²; and sensus of attention represents the level of conscious arousal (or ‘wakefulness’ [100]) 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.

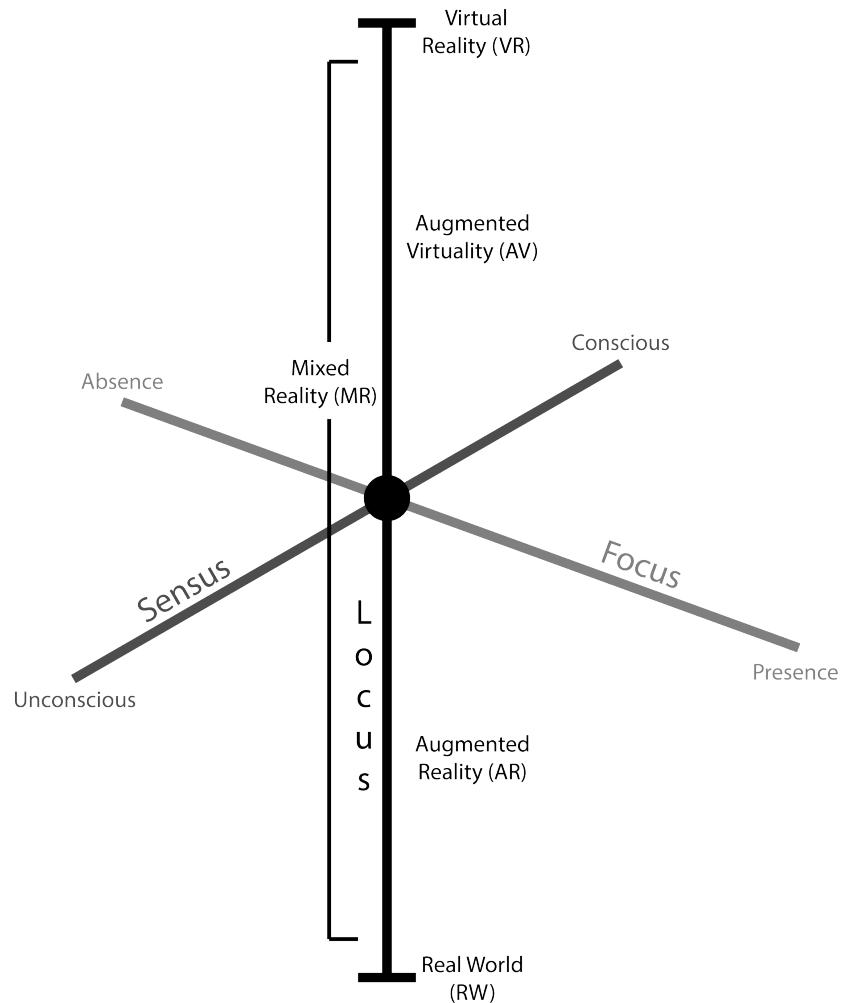


Figure 5.6: The combined virtuality continuum/three dimensions of virtual experience model.

¹ **Presence** in this context is defined as a state of heightened perceptual processing of environmental stimuli (“*a psychological focus on direct perceptual processing*” [99]) 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*” [99].

5.1.3 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.

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 [101];
- there is existing VR content;
- the visual differences between real & virtual environments are so substantial that an AR system would resort to augment (&/or diminish [102]) almost the whole RW view. While AR "*smears an informational coating over real space*" [103], 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.

5.1.4 Example Application - Cultural Heritage

The field of cultural heritage has seen widespread applications of both AR [104–118] & VR [101, 119–124]. 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).

5.1.5 Case Study - St Salvator's Chapel

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 5.7 shows the 1450-1460 layout of the chapel (including the paths that the IPS has been prepared upon).

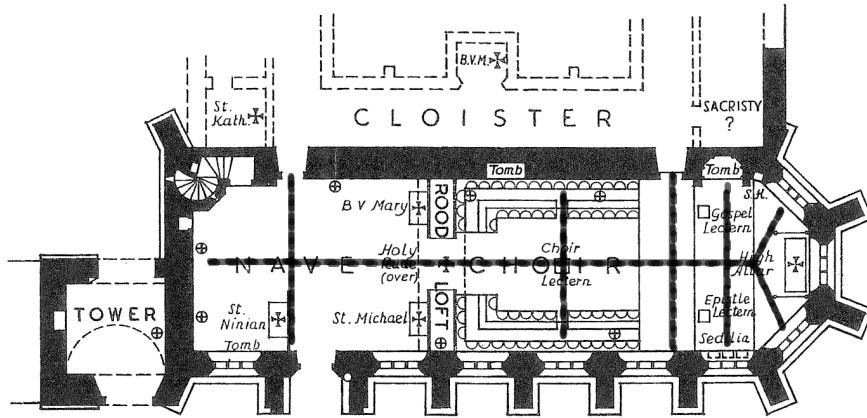


Figure 5.7: Floor plan of St Salvator's chapel, with IPS routes.

5.2 The Mirrorshades Platform

Figure 5.8 presents a high level architectural overview of our mobile XR platform, dubbed Mirrorshades³.

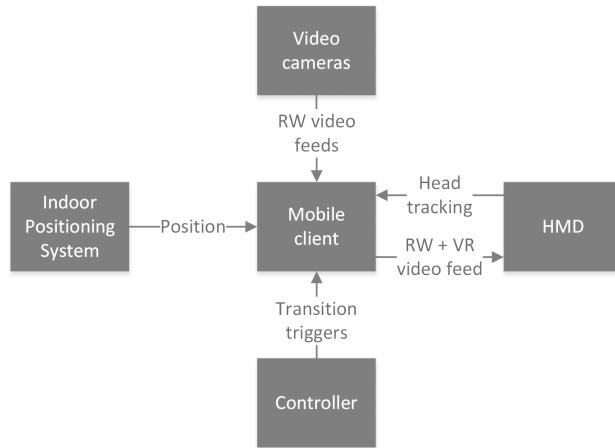


Figure 5.8: Overview of the Mirrorshades platform.

5.2.1 Implementation

Figure 5.9 presents an overview of the implementation of the Mirrorshades platform design for use in the chapel investigations.

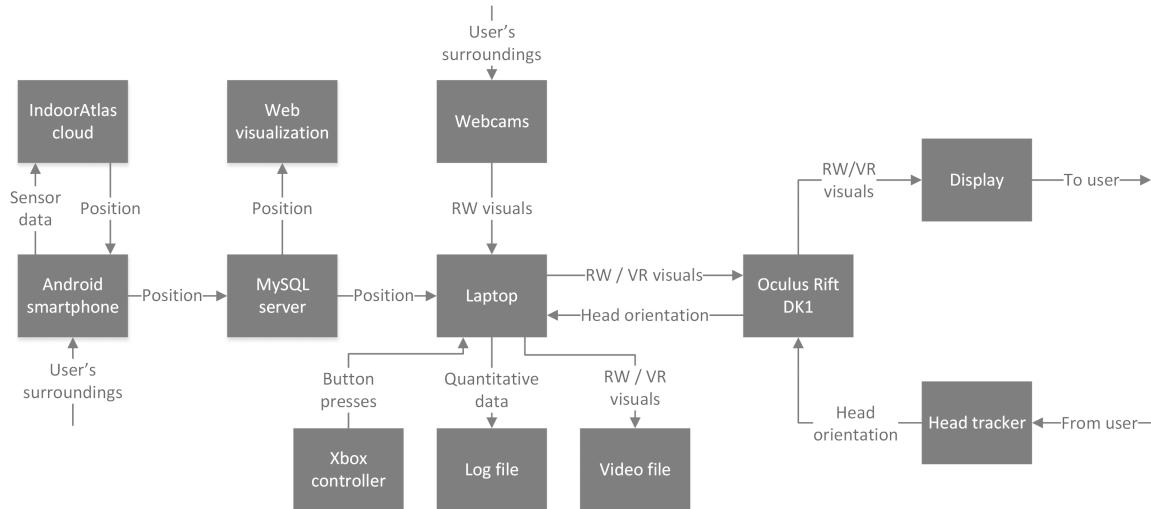


Figure 5.9: Implementation of Mirrorshades platform.

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

5.2.2 Hardware Components

The hardware of the implementation comprises;

- 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 5.10);
- 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 5.10: HMD with stereo camera solution.

5.2.3 Software Components

The software of the implementation comprises;

- an Android application that runs on the smartphone, determines the location of the phone within the building that it is in using the IndoorAtlas IPS [125] (figure 5.7 shows the paths within the chapel upon which the IPS has been configured) & submits these location data via PHP to a database server;
- a MySQL database server that stores location data for the phone & allows these data to be accessed both by the Unity application running upon the laptop & by a web visualisation;
- a Unity application that runs on the laptop.

5.2.4 Integration of Components

The Unity application hosts the VR representation of the chapel & takes in feeds from both webcams, the HMD head tracker & the 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.

5.3 Investigation 1 - The Case for Mobile XR

This first investigation will compare interaction with the RW & VR chapel using Mirrorshades to interaction with the same content separately, the latter being the approach usually adopted for dissemination of VR content in cultural heritage contexts. Participants will complete a task that will promote active comparison & contrast of the RW & VR environments, whilst navigating a set route. This investigation will gauge through experimentation whether the Mirrorshades platform provides any value over the traditional manner in which the same VR content might be disseminated at a cultural heritage site.

5.3.1 Setting & Task

This investigation comprises two phases;

1. Participants will experience the RW & VR chapels separately. They will navigate the VR chapel from a stationary position, as one might expect to see a VR installation at a cultural heritage site, using the Xbox controller to move around the VR environment observed via the HMD. The HMD will obscure their view of the RW chapel around them. Subsequently, they will navigate the RW chapel without the HMD or any associated equipment.
2. Participants will experience the RW & VR chapels in tandem using the Mirrorshades platform. They will wear the HMD, holding the Xbox controller in their right hand & the smartphone in their left, with the laptop & battery pack in a satchel worn over one shoulder. One style of transition will be available to the participants during this phase.

In all 3 scenarios (phase 1 VR, phase 1 RW, phase 2 RW + VR) the participant will navigate the same route & will be instructed to identify a particular feature/object within the chapel (see figures ?? & ??), situated somewhere upon the set route, that differs in its appearance &/or location between the RW & VR chapels. The order in which the two phases are completed will be randomised between participants, as will the features that they are told to observe in each phase. Participants will have a maximum length of time to navigate the route & will be allowed to stop before this time has elapsed should they wish.

5.3.2 Evaluation Techniques

Evaluation will be performed via a short structured interview & completion of the System Usability Scale (SUS) [126].

5.3.3 Hypothesis

SUS responses are expected to average fairly low due to the cumbersome nature of the platform's implementation. Participants who are able to overcome this cumbersomeness are expected to respond favourably to the platform, with those who cannot overcome it responding in favour of the traditional 'separate' approach instead.

5.4 Investigation 2 - Informing Mobile XR Implementation

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.

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 5.1.2) 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 [127], 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) [128] & that when engaging with VR content a user's focus can even be said to typically be *shared* between VR & RW [99].

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 [129, 130] 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 [131].

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

Transitions can be performed in multiple different manners & it is hypothesized that users will prefer different styles of transition in different situations, surroundings & scenarios (where 'preference' toward a particular style of transition is expected to correlate strongly with a less severe BIP being experienced upon its execution).

To this end, several different transition methods have been developed & this investigation will endeavour to identify & quantify preferences toward them, to infer which approaches to transitioning between RW

⁴The definition of **break in presence** adopted herein is the second from Waterworth & Waterworth [99] (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 [132] as they considered presence only in terms of attending to stimuli from a virtual environment, with a break in presence 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'.

& 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 will be a strong correlation between participant movement (or lack thereof) & choice of particular transition style.

5.4.1 Transitions using the Combined Model

Visualised using the combined model (see section 5.1.2) as figure 5.11, 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

5.4.2 Transition Methods

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 5.12 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.

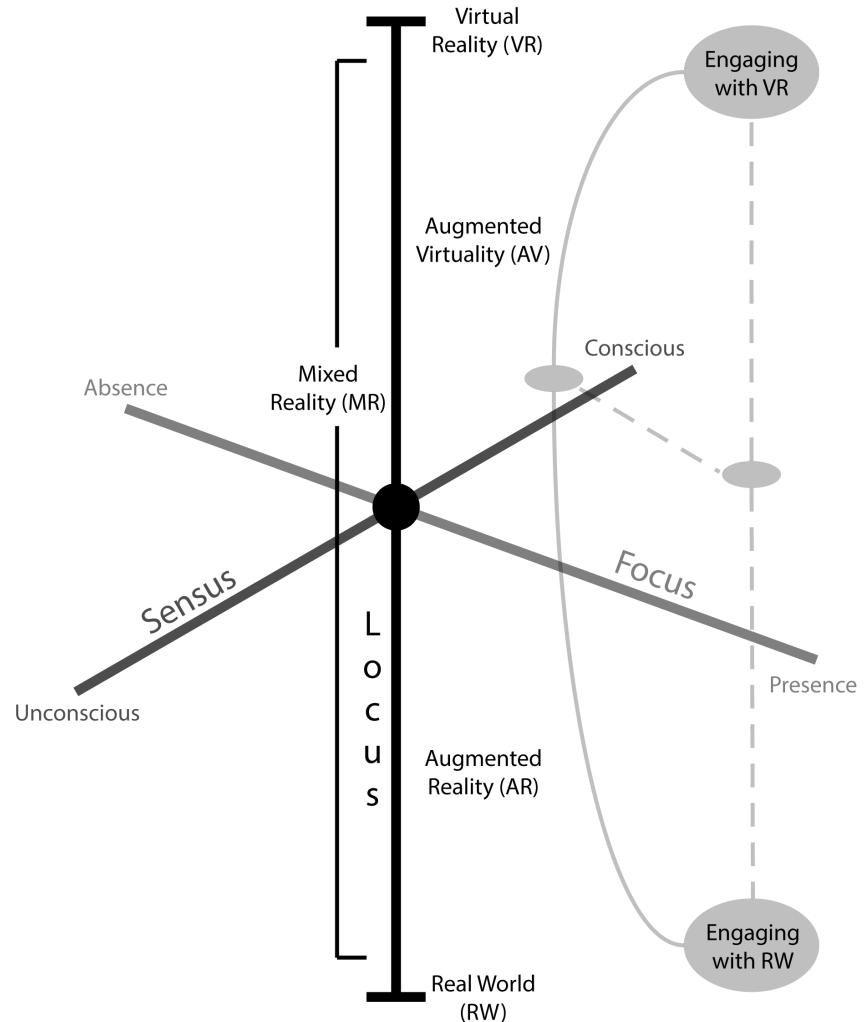


Figure 5.11: Operation of the Mirrorshades platform represented upon the combined model.

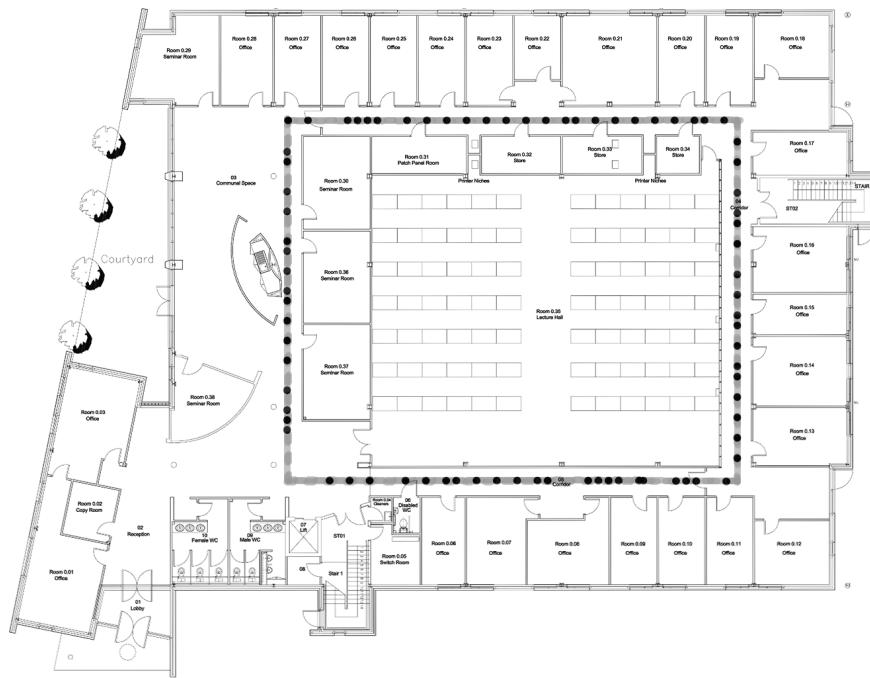


Figure 5.12: Positions (black circles) reported whilst walking a slow lap ($< 1\text{ms}^{-1}$, following gray path) of a departmental building. The building is approximately 40m wide by 30m tall.

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 5.13 illustrates this scenario.

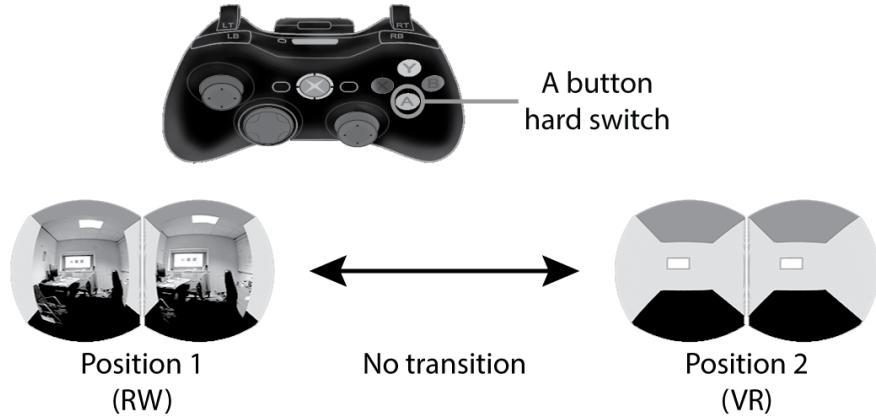


Figure 5.13: Hard switch.

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 5.14 illustrates this scenario.



Figure 5.14: Switch with linear interpolation.

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 5.15 illustrates this scenario.

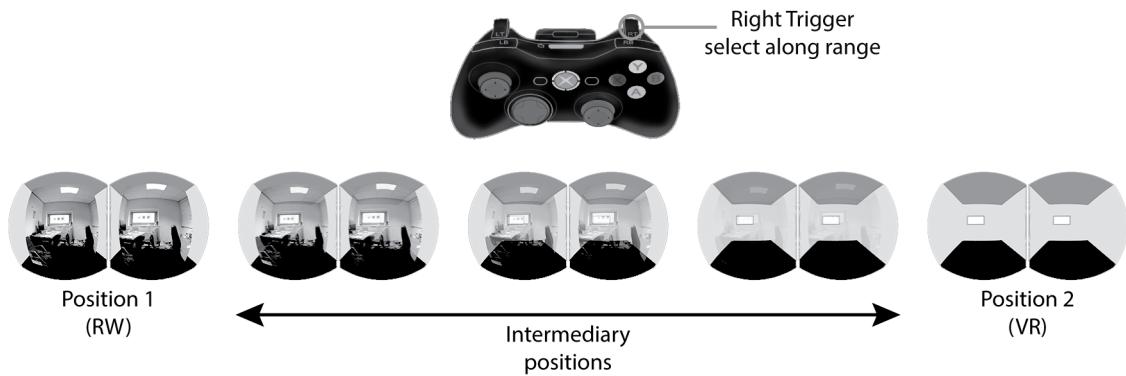


Figure 5.15: Analogue selectable opacity.

Periodic hard switches

Independent or in addition to any of the previous scenarios, the visual stimuli displayed by the HMD switch from RW to VR at a set interval & for a set amount of time. For example, every 3 seconds the stimuli switch from RW to VR for 0.2 of a second before switching back from VR to RW. Any user triggered transitions cause the interval timer to be reset, such that an 'automated' switch will never occur after less time from a user triggered switch than the set interval. Automated transitions are disabled whilst [RT] is at all depressed. Figure 5.16 illustrates this scenario, where i represents the interval between switches & d represents the duration of the switch from RW to VR.

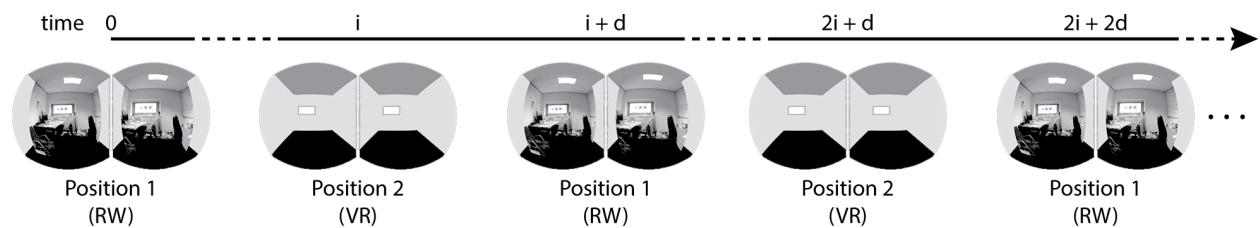


Figure 5.16: Periodic hard switches.

Reduced maximum opacity

Independent or in addition to any of the previous scenarios, the maximum opacity of the game objects that the webcam feeds are rendered upon is reduced, such that the ‘default’ position at which a transition has not been triggered (either by a button press, trigger movement or by a periodic switch) displays VR superimposed upon RW. Figure 5.17 illustrates this scenario in combination with a hard switch (from section 5.4.2) in which the user triggers hard switches between the default position of a superimposition of VR upon RW & a position where only VR stimuli are present.

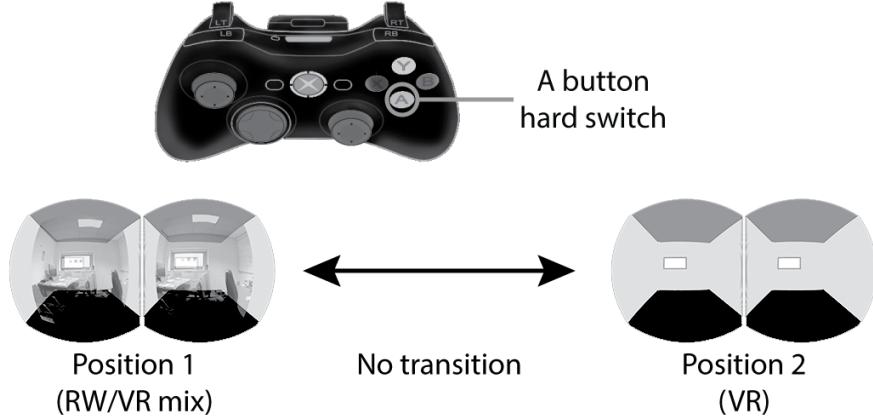


Figure 5.17: Hard switch from reduced maximum opacity.

5.4.3 Experimental Task & Setting

For these experiments, the HMD is worn upon the head of the participant & is connected to the laptop computer, 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).

A task similar to that employed in the first investigation (see section 5.3 will be employed, encouraging participants to encounter multiple different scenarios of moving, remaining stationary, etc.

5.4.4 Evaluation

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 [128]. Thus, subjective measures will produce the bulk of the data for evaluation. However, objective data will also be 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 [99], one of the subjective measures that will be used in this evaluation will be an established presence measure, to try to capture the behaviour of the user’s position upon the focus axis.

Subjective Quantitative - Post-Task Questionnaire

After completing the task, participants will respond to the Igroup Presence Questionnaire (IPQ) [133] (see appendix 5.5 for the items of the IPQ) which will provide subjective quantitative insight into their experiences

with the system, in particular in relation to their position upon the focus axis of the combined model. The IPQ represents a useful questionnaire for evaluation of users' subjective experiences of using the Mirrorshades platform because its terms, especially in the 'spatial involvement' scale, question about the RW environment in a manner that does not explicitly present it as a 'distraction' from the VR interaction as many other presence questionnaires do.

Whilst a traditional VR experience would hope to elicit high SP1 & SP4 results combined with low INV1 & INV3 results, Mirrorshades participants are expected to report high SP1 & SP4 combined with *high* INV1 & INV3. The results from participants in this investigation will be compared against those who partook in a 'traditional' VR experience wherein RW stimuli were considered a distraction.

Subjective Qualitative - Interview

A structured interview will be performed after the IPQ has been completed.

Objective Quantitative - Automatic Data Logging

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

- <frame_number>
- <timestep> - 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_opacity> - the maximum opacity of the game objects upon which the webcam feeds are rendered (see section 5.4.2)
- <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 5.4.2)
- <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 are expected to 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.

Objective Qualitative - Video Recording

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.

5.4.5 Preliminary Results

Figure 5.18 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 switching to a semi opaque virtual feed rather than complete transitions to 100% virtual which appear to be more common when standing still; for example, frames 500-700, 3000-3300 & 4000-4100.

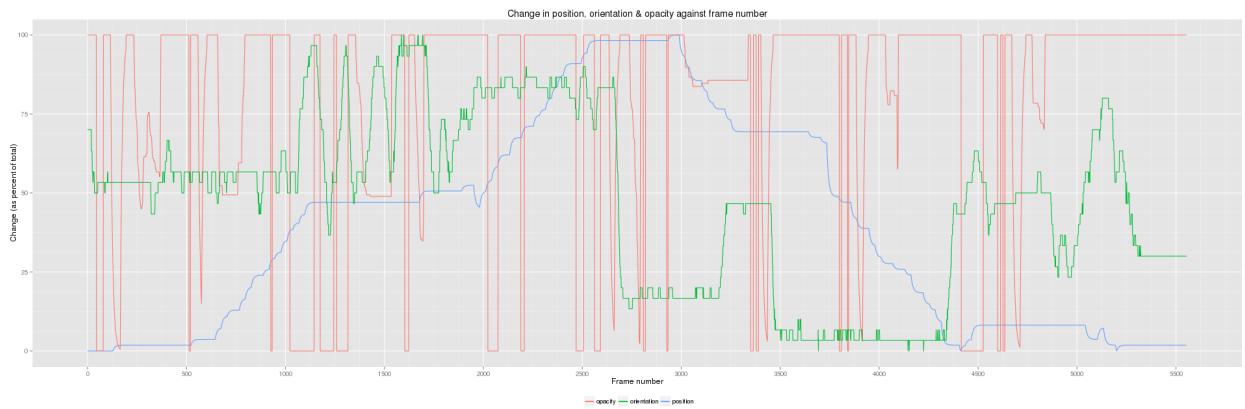


Figure 5.18: Plot of a subset of example objective quantitative data logged by the Unity application.

5.5 IPQ Items

G1: In the computer generated world I had a sense of “being there”.

SP1: Somehow I felt that the virtual world surrounded me.

SP2: I felt like I was just perceiving pictures.

SP3: I did not feel present in the virtual space.

SP4: I had a sense of acting in the virtual space, rather than operating something from outside.

SP5: I felt present in the virtual space.

INV1: How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

INV2: I was not aware of my real environment.

INV3: I still paid attention to the real environment.

INV4: I was completely captivated by the virtual world.

REAL1: How real did the virtual world seem to you?

REAL2: How much did your experience in the virtual environment seem consistent with your real world experience?

REAL3: How real did the virtual world seem to you?

REAL4: The virtual world seemed more realistic than the real world.

6

Mirrorshades - Studies/Results

The Case for Mobile XR

This investigation compares two scenarios for interaction with a real location & a corresponding virtual location.

1. **Stationary scenario** - interacting with the virtual location from a fixed real location, then subsequently interacting with the real location.
2. **Mobile scenario** - using the Mirrorshades platform to interact with both the real location & the corresponding virtual location in tandem, whilst moving around both environments.

The locations in question are St Salvator's chapel & a virtual reconstruction of the chapel as it stood in 1450-1460. The stationary scenario is representative of how virtual reality technologies, including both CAVEs & HMDs, have previously been used for dissemination of virtual reality content in cultural heritage contexts [121] & thus this investigation serves to compare Mirrorshades with previous applications virtual reality content to these contexts.

Process

- Participants complete a pre-task questionnaire, which provides calibration for their subsequent responses by enquiring about age, gender identity, previous experience with VR hardware & previous interactions with either the real or virtual chapel. This questionnaire is included as Appendix 6.
- Participants familiarise themselves with the experience of using the Oculus Rift DK1 HMD & the Xbox 360 controller by interacting with the 'Tuscany demo' prepared & maintained by the Oculus VR team. This is performed from a seated position.
- Participants complete the stationary scenario.
- After completing the stationary scenario, participants complete the System Usability Scale (SUS) [126] questionnaire, included as Appendix 6 & a 12-item questionnaire, included as Appendix 6.
- Participants complete the mobile scenario.
- After completing the mobile scenario, participants complete the SUS questionnaire & the 12-item questionnaire again.
- Finally, the participant is engaged in a short structured interview. Interview prompts are included as Appendix 6.

In addition to SUS, the 12-item questionnaire & the structured interview, quantitative data is logged by the Mirrorshades platform when a participant is interacting with virtual content in the first scenario & at all times during the second scenario.

The Scenarios

Both scenarios that participants complete for this investigation are designed to mimic the style of exploration & interaction that visitors to the chapel display, which was observed over several occasions. Visitors enter the chapel from the North/West corner then proceed to walk Eastwards along the nave, pausing to look around after passing through the rood screen, before continuing along the nave toward the altar. Visitors pause in front of the alter upon reaching the end of the pews & then walk North toward the tomb where they pause again to inspect it. Participants are instructed to imagine that they are performing a similar visit to the chapel & to follow a similar path, pausing after the rood screen, at the end of the pews & in front of the tomb. Participants are shown the map included as figure 6.1 to explain the scenario better.

In the stationary scenario, participants interact with the virtual chapel using the Rift & Xbox controller, whilst in a sitting position. After completing the path, they remove the headset & then walk the same path in the real chapel. This scenario alludes to how virtual reality technologies have previously been applied to cultural heritage situations, allowing visitors to experience a virtual reality reconstruction or reimagination of the real environment from a fixed position & with their view of the real environment wholly occluded by their view of the virtual environment.

In the mobile scenario, participants wear the HMD, hold the Xbox controller in their right hand & the smartphone in their left, with the laptop & battery pack in a satchel worn over the right shoulder. They then walk the same path, but this time with the ability to transition at any time between viewing the real environment & the virtual environment from the same vantage point.

In this first investigation phase, only one transition is available to participants. Preliminary experiments involving the researchers' colleagues that allowed hard transitions, linear interpolated transitions & analogue selectable opacity, indicated that the linear interpolated transition was preferred to either the hard transition or the analogue selectable opacity & thus this is the transition available to participants in this first phase investigation.

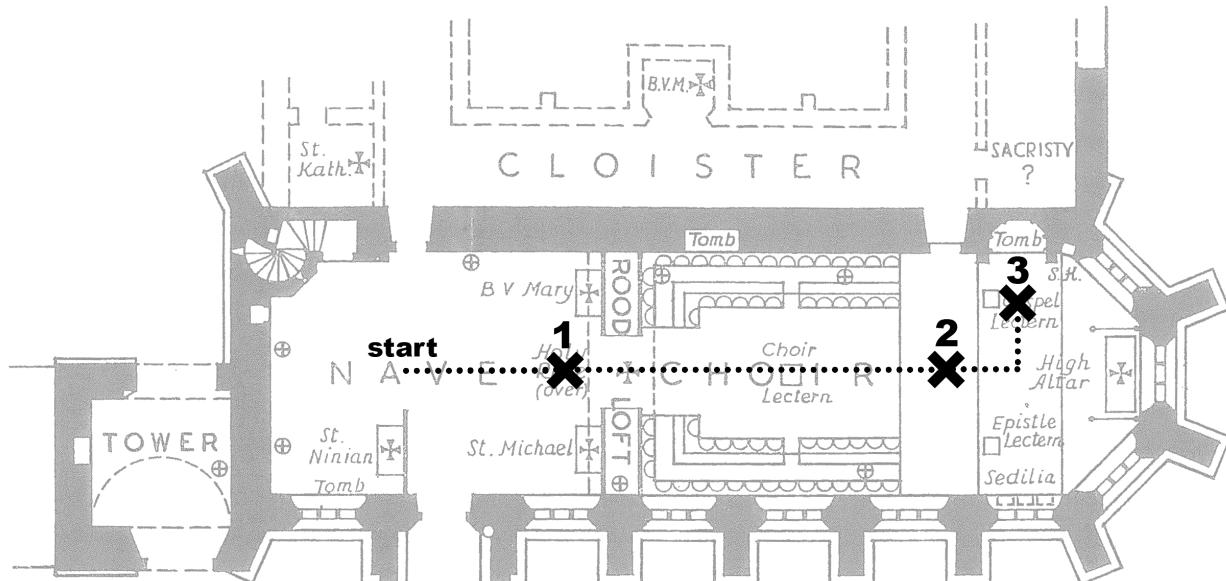


Figure 6.1: The path & positions within the chapel that participants are instructed to attend to.

Hypotheses

The aim of the mobile scenario is to improve participant engagement with & understanding of the relationships between the real & virtual environments, by addressing the problems of spatial & temporal separation inherent with the ‘traditional’ stationary scenario, by imparting upon the participant the ability to transition between equivalent vantage points within the real & virtual environments at will.

While we expect participants to report that the mobile scenario does indeed allow them to better compare & contrast the real & virtual environments, identify differences between the real & virtual environments & gain a better understanding of how the real & virtual environments relate to each other, we expect some participants to report that having to ‘split’ their attention between the two environments in the mobile scenario leads to lessened engagement & understanding & that the visual quality of the real view through the headset/cameras leads to some participants preferring to interact with the real environment without the headset.

We expect the cumbersome nature of the mobile scenario & the reduced quality of viewing the real environment via the headset/cameras to have a noticeable effect upon participants movement (both position & head orientation) in the mobile scenario.

Addressing these issues, such that participants don’t find viewing the real through the headset to be such a reduction in quality compared to just seeing real, such that participants feel as though they can move & look around themselves as much in the mobile scenario as in the stationary scenario & such that participants transition between real & virtual at any time instead of avoiding transitions in situations in which they think that they will be unpleasant/jarring, is key & what the next stage will focus on.

SUS

SUS scores for the mobile scenario are expected to average lower than those for the stationary scenario, due to the cumbersome nature of the platform when performing the mobile scenario; during the stationary scenario, participants are seated, whilst during the mobile scenario they are required to carry a satchel over one shoulder & hold a smartphone in their left hand. Participants who are able to overcome this cumbersomeness are expected to respond more favourably to the mobile scenario than those who cannot overcome it.

12-item Questionnaire

- Participants will find it easier to compare & contrast real & virtual environments in the mobile scenario than in the stationary scenario (q2)
- Participants will experience a greater sense of ‘being in’ the virtual environment in the mobile scenario than in the stationary scenario (q4, due to physical movement/embodiment)
- Participants will have a greater sense of ‘being in the past’ in the mobile scenario than in the stationary scenario (q7)
- Participants will maintain greater awareness of both real & virtual environments in the mobile scenario than in the stationary scenario (q5)
- Participants will gain a better understanding of what the chapel was like in the past in the mobile scenario than in the stationary scenario (q12)

Log data

- Head movement (pitch & yaw) will be more restricted in the mobile scenario compared to the stationary scenario
- Aversion to looking around (even at real) when moving in the mobile scenario

- Head movements will be larger discrete changes in the stationary scenario compared to the mobile scenario
- Tendency to only look at virtual when looking around

Interviews

- mobile scenario makes it easier to spot differences
- mobile scenario reveals differences that stationary didn't
- stationary does not reveal differences that mobile doesn't
- mobile scenario is preferred & is user-reported as 'more engaging'

Results

Pre-task Questionnaire

For n=5 ages ranged from 21-26, 3x female & 2x male, all reported previous experience using a games console controller, 1x reported previous use of a HMD, 2x reported having previously visited the chapel, none had previously interacted with the virtual chapel model.

SUS

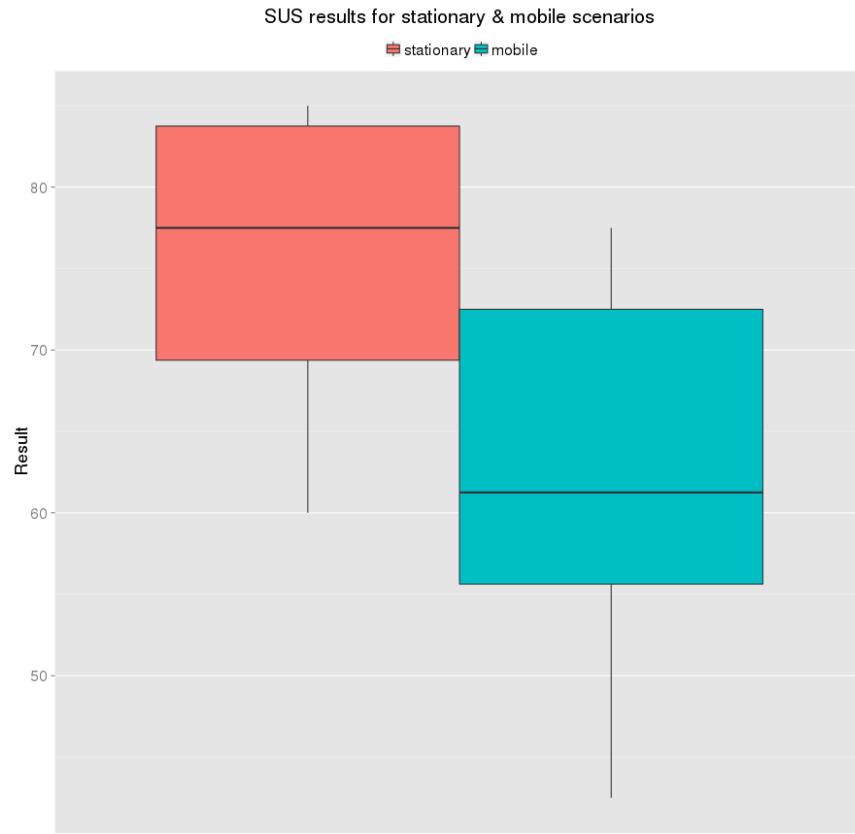


Figure 6.2: SUS results.

As expected, the SUS scores for the mobile scenario are lower than those of the stationary scenario, although not drastically so. Furthermore, although scoring lower on SUS, the mobile scenario came out above the stationary scenario when looking at the results of q8 in the 12-item questionnaire which asked participants if they thought they would have preferred a conventional computer monitor.

12-item Questionnaire

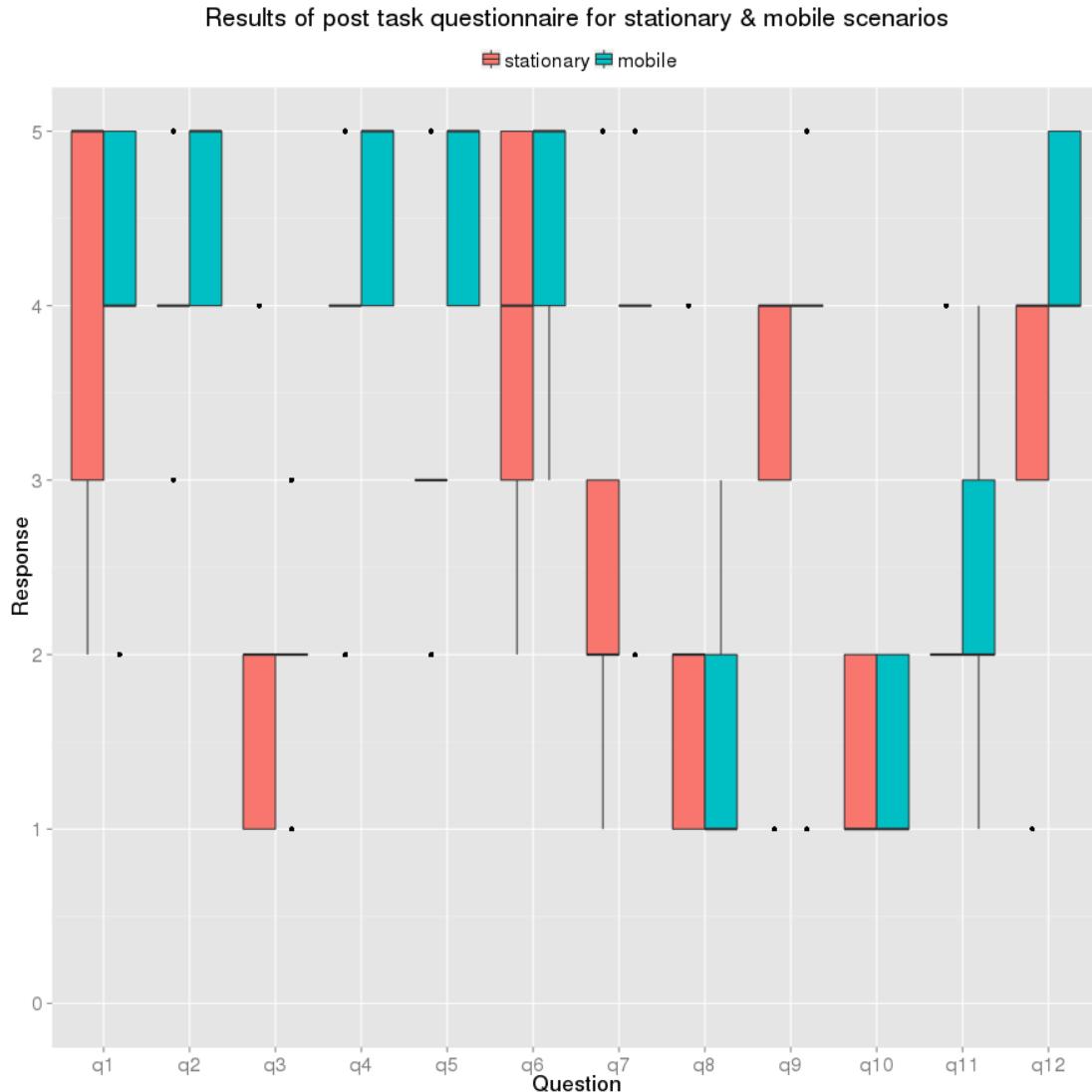


Figure 6.3: 12-item questionnaire results.

The hypotheses seem to hold, in particular;

- Participants will maintain greater awareness of both real & virtual environments in the mobile scenario than in the stationary scenario is supported by the responses to q5
- Participants will have a greater sense of ‘being in the past’ in the mobile scenario than in the stationary scenario is supported by q7 (thanks to embodiment?)
- Participants will gain a better understanding of what the chapel was like in the past in the mobile scenario than in the stationary scenario is supported by q12

It is worth highlighting the responses to q10 in relation to those to q2. Participants reported finding it easier to compare features from the past & present (q2) during the mobile scenario, however did not report a difference between not noticing differences between the real & virtual environments (q10).

Log data

***stats here (see R output files)

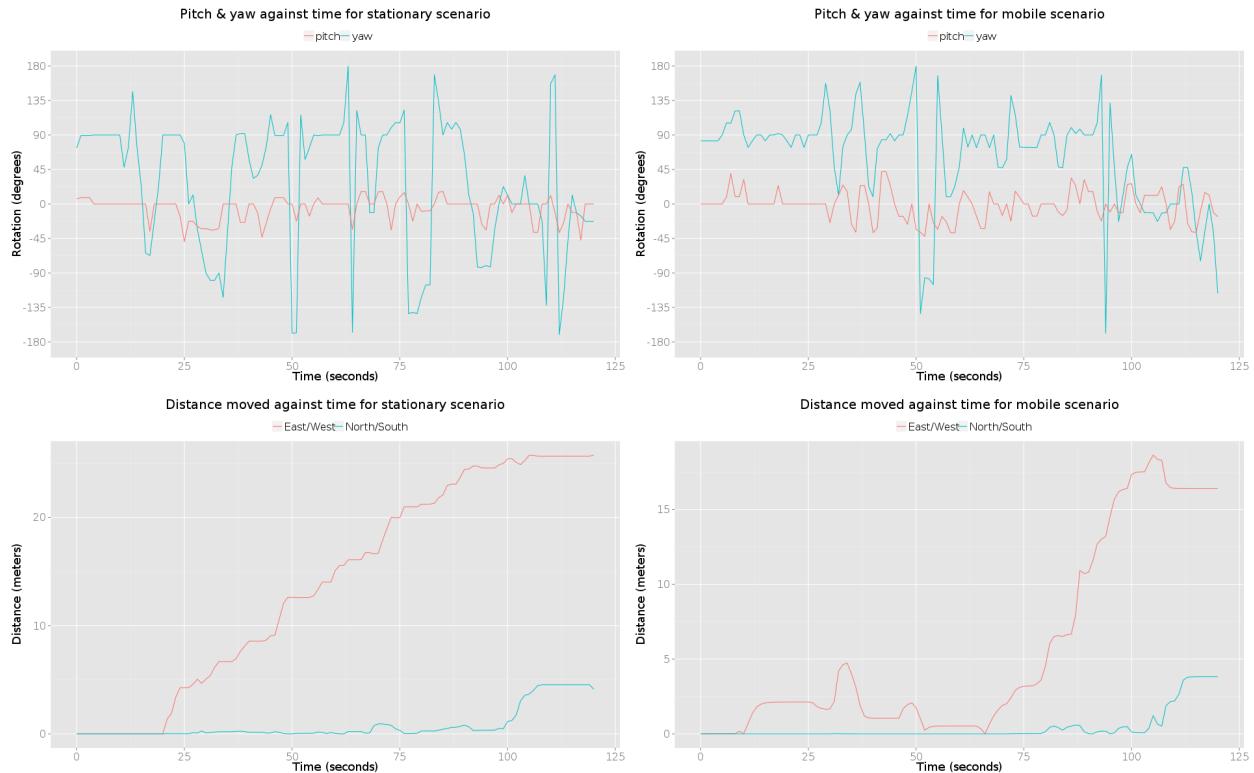


Figure 6.4: Participant 1

Interviews

Pre Task Questionnaire

Post Task Questionnaire (SUS)

Post Task Questionnaire (Own)

Structured Interview Questions

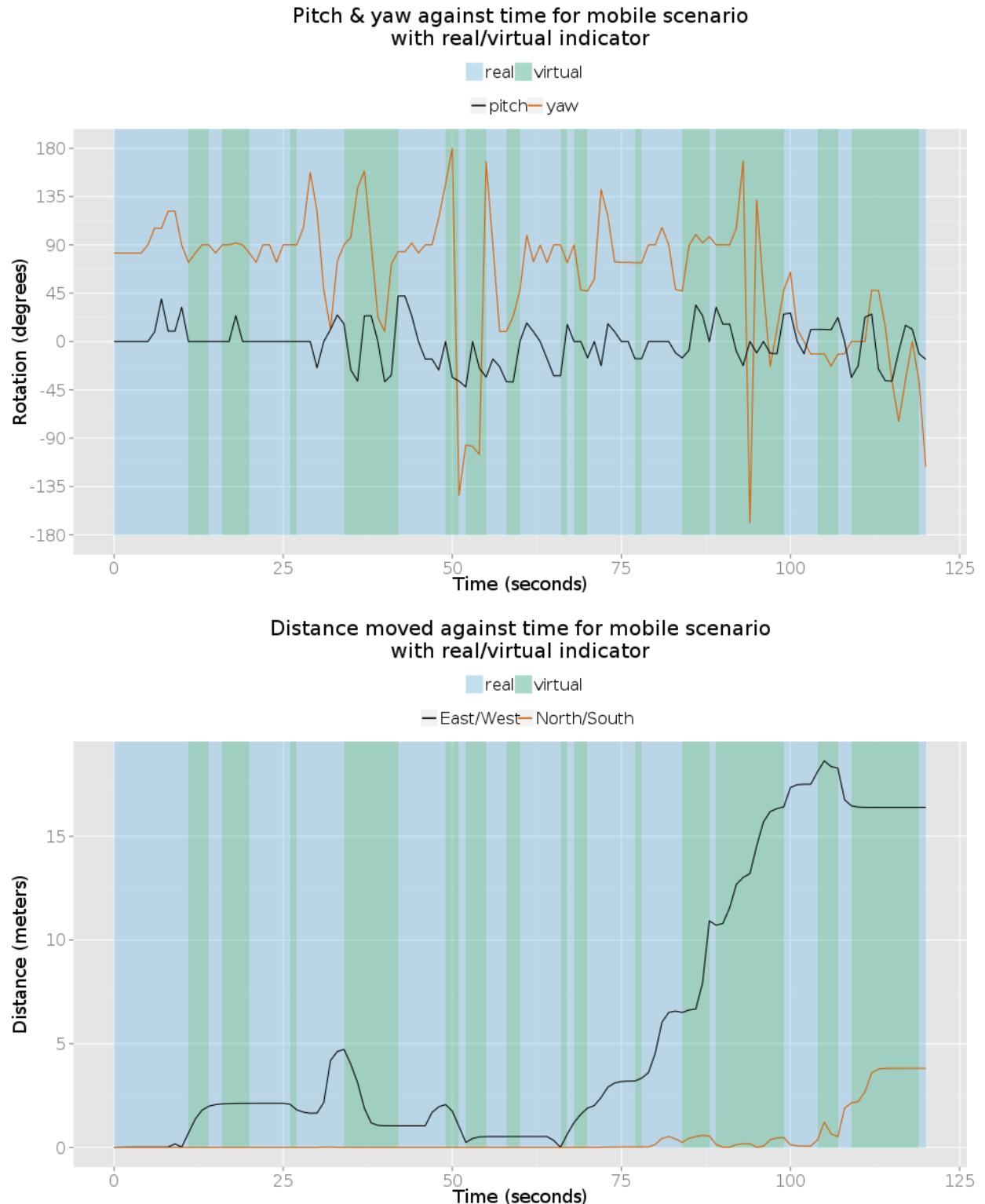


Figure 6.5: Participant 1

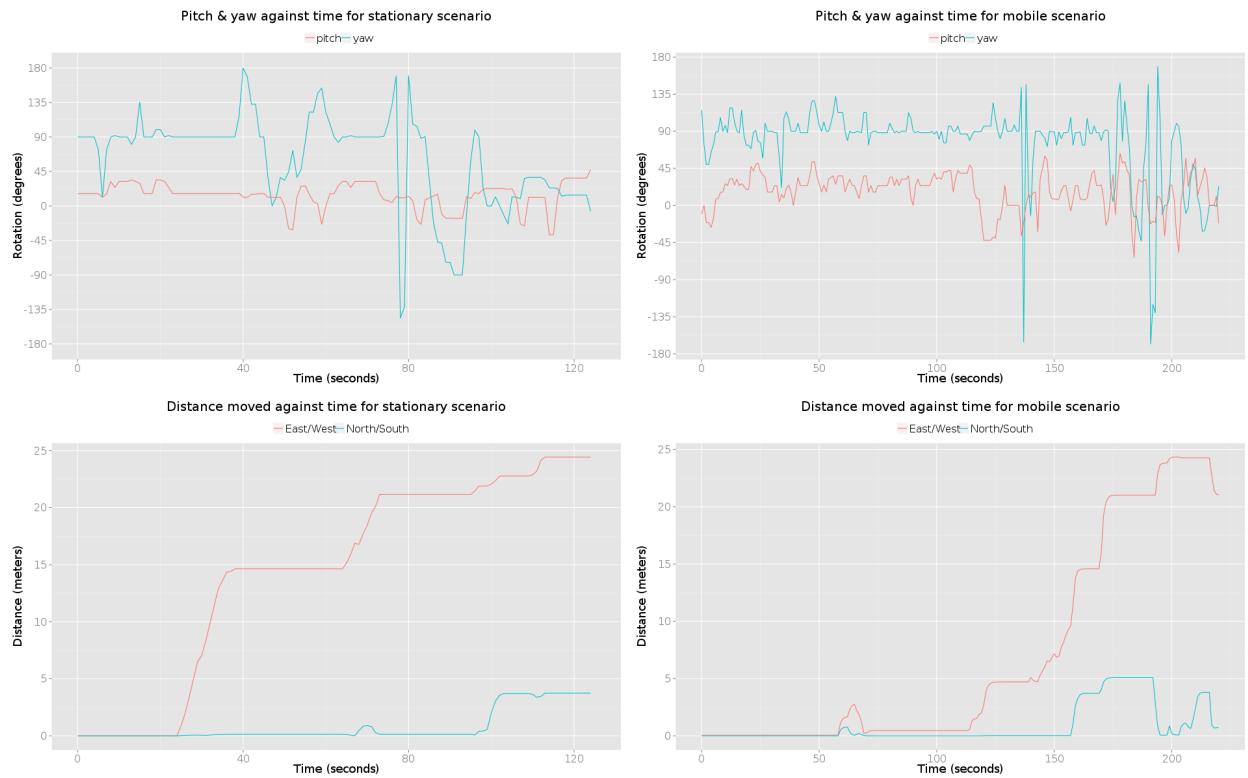


Figure 6.6: Participant 2

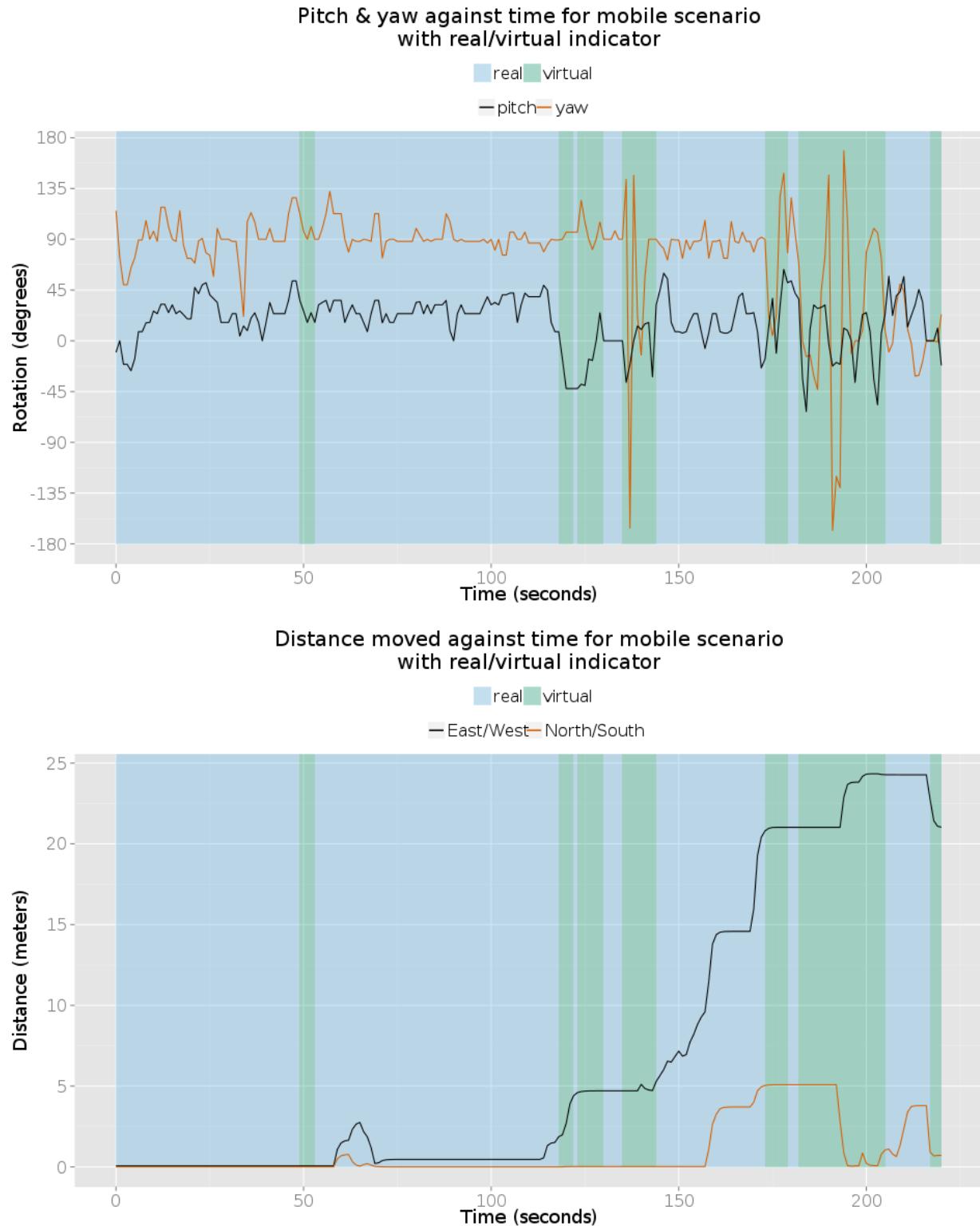


Figure 6.7: Participant 2

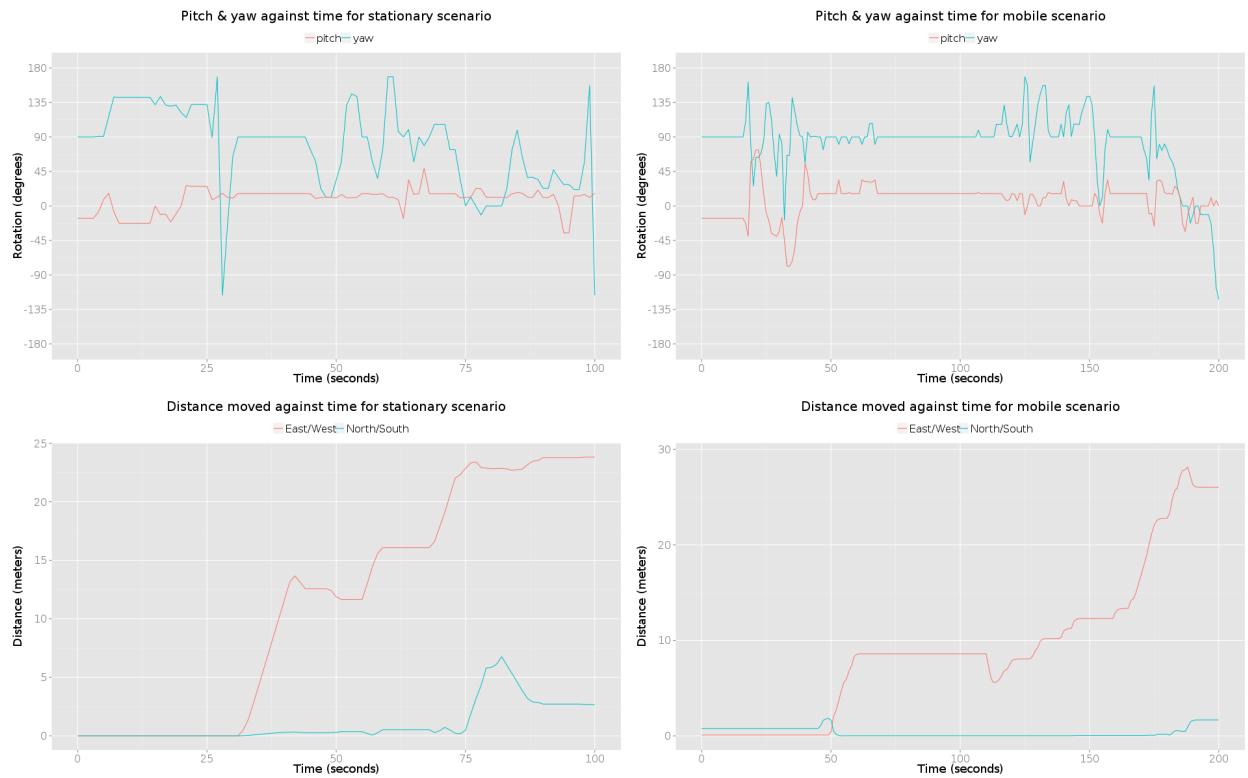


Figure 6.8: Participant 3

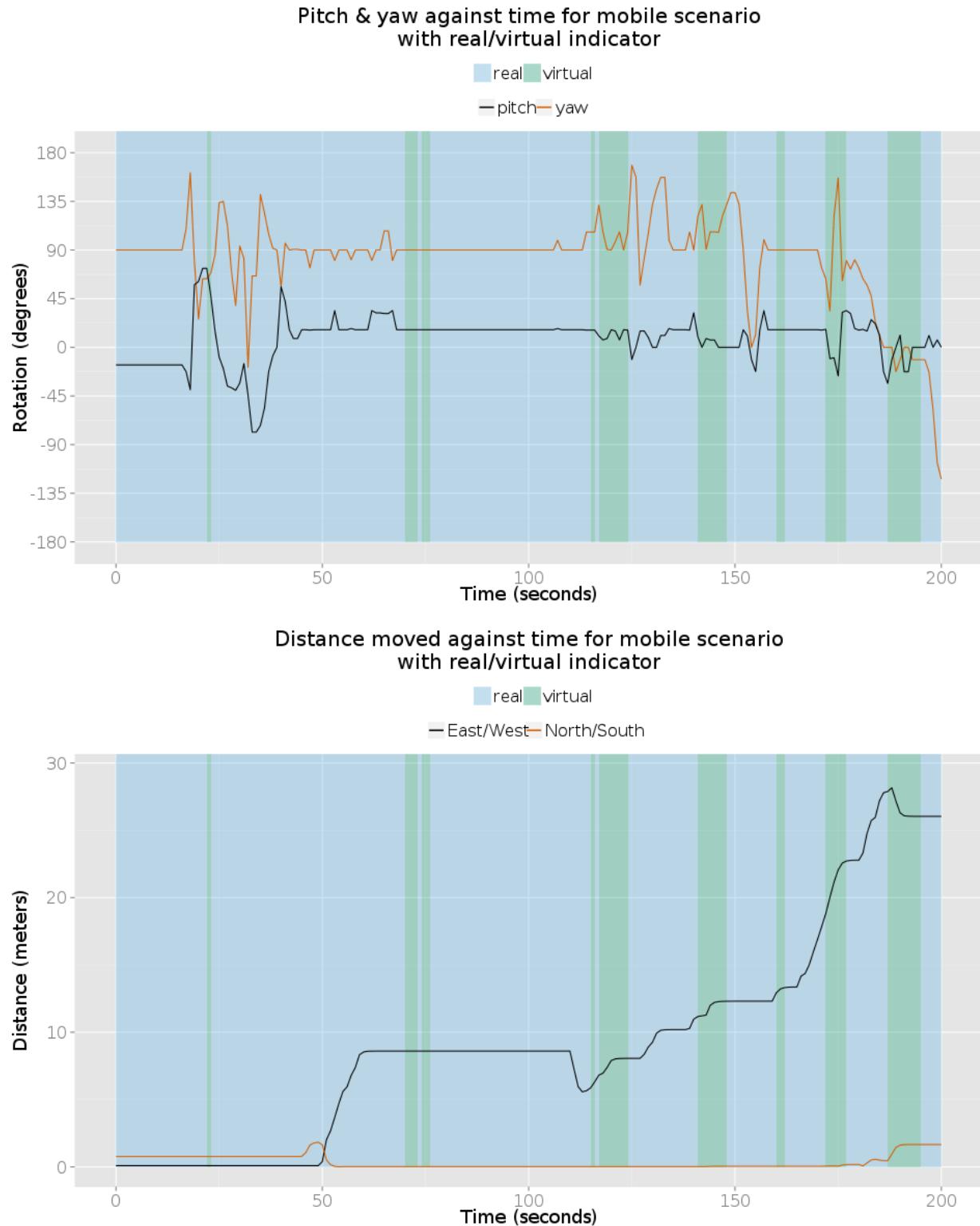


Figure 6.9: Participant 3

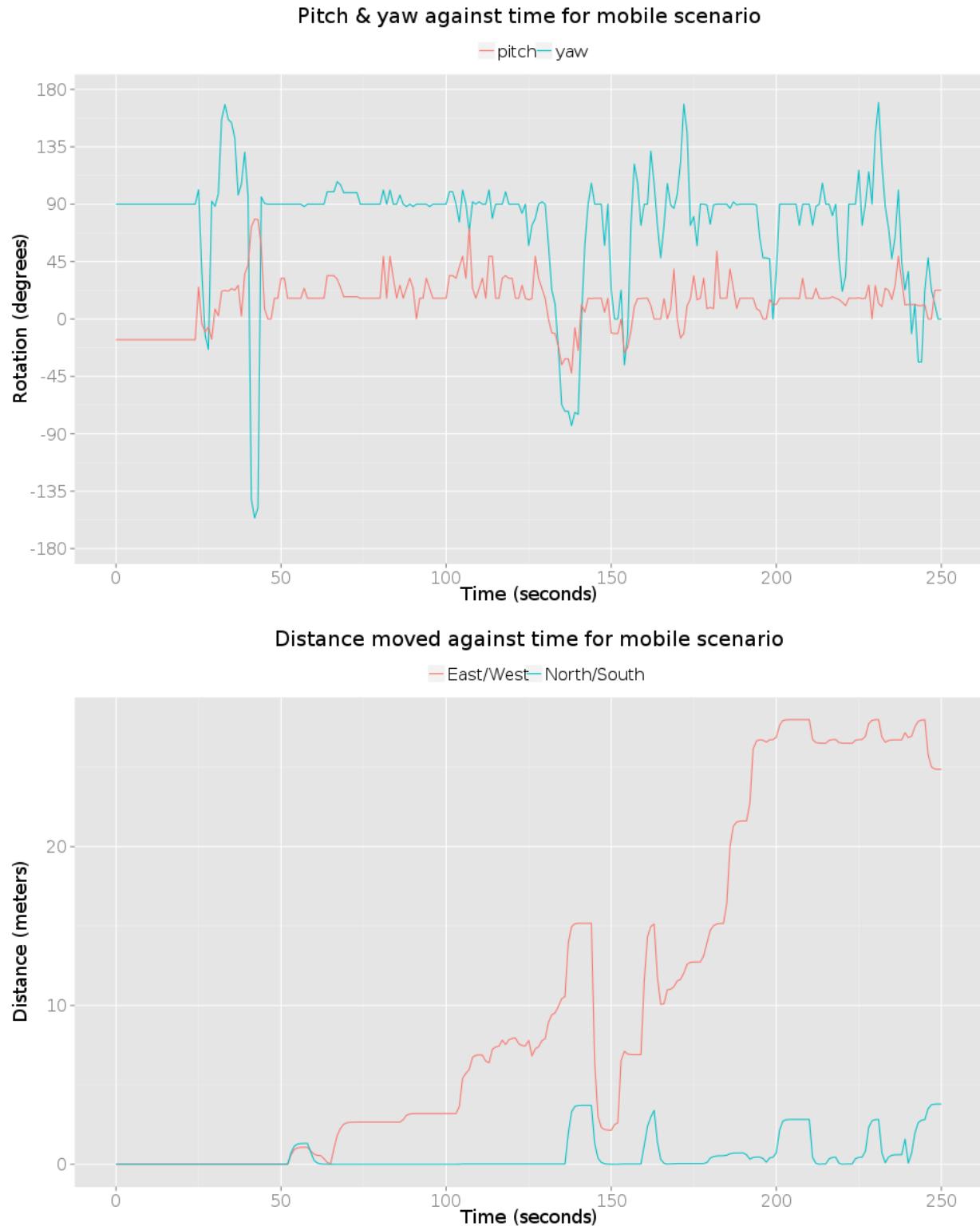


Figure 6.10: Participant 4

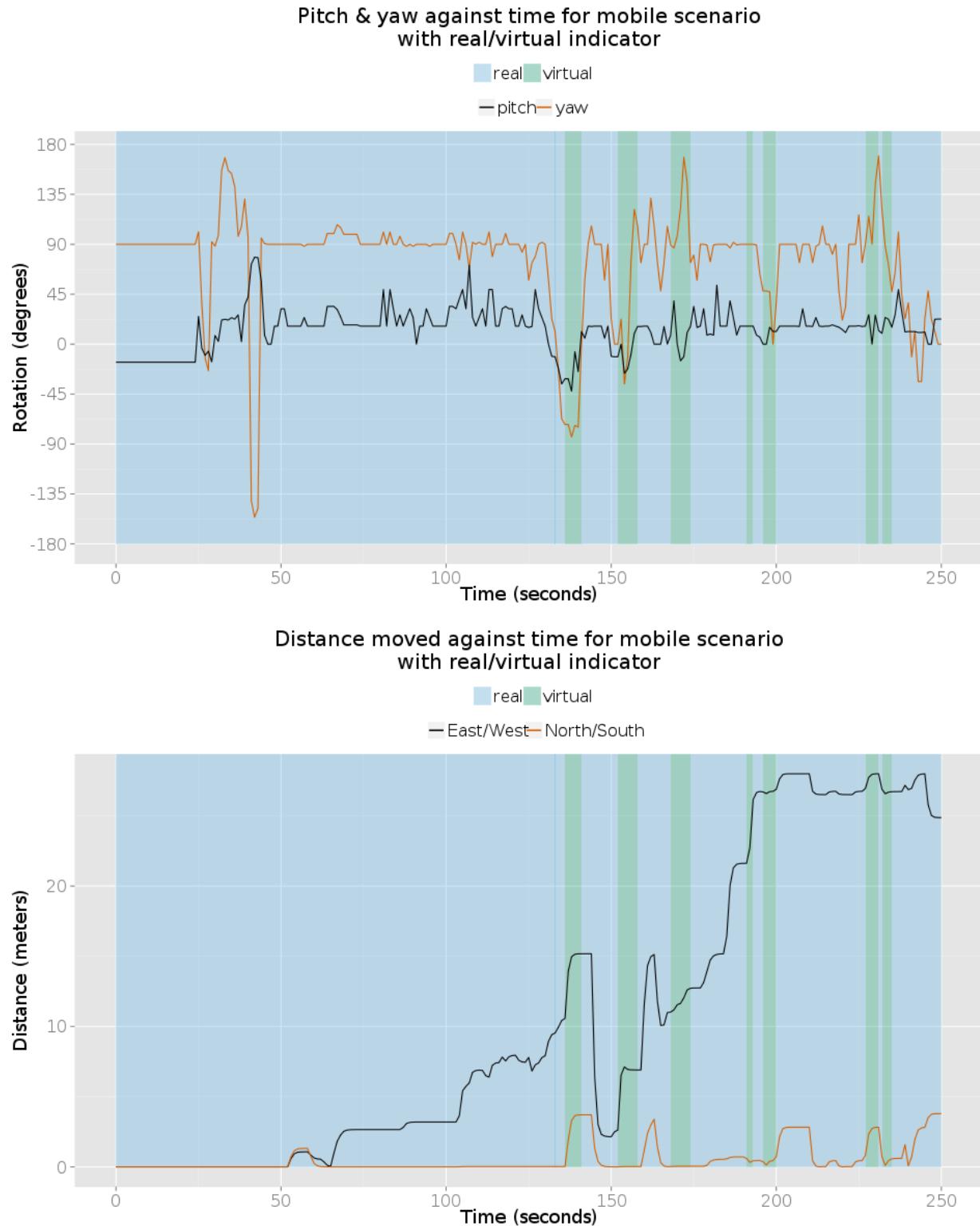


Figure 6.11: Participant 4

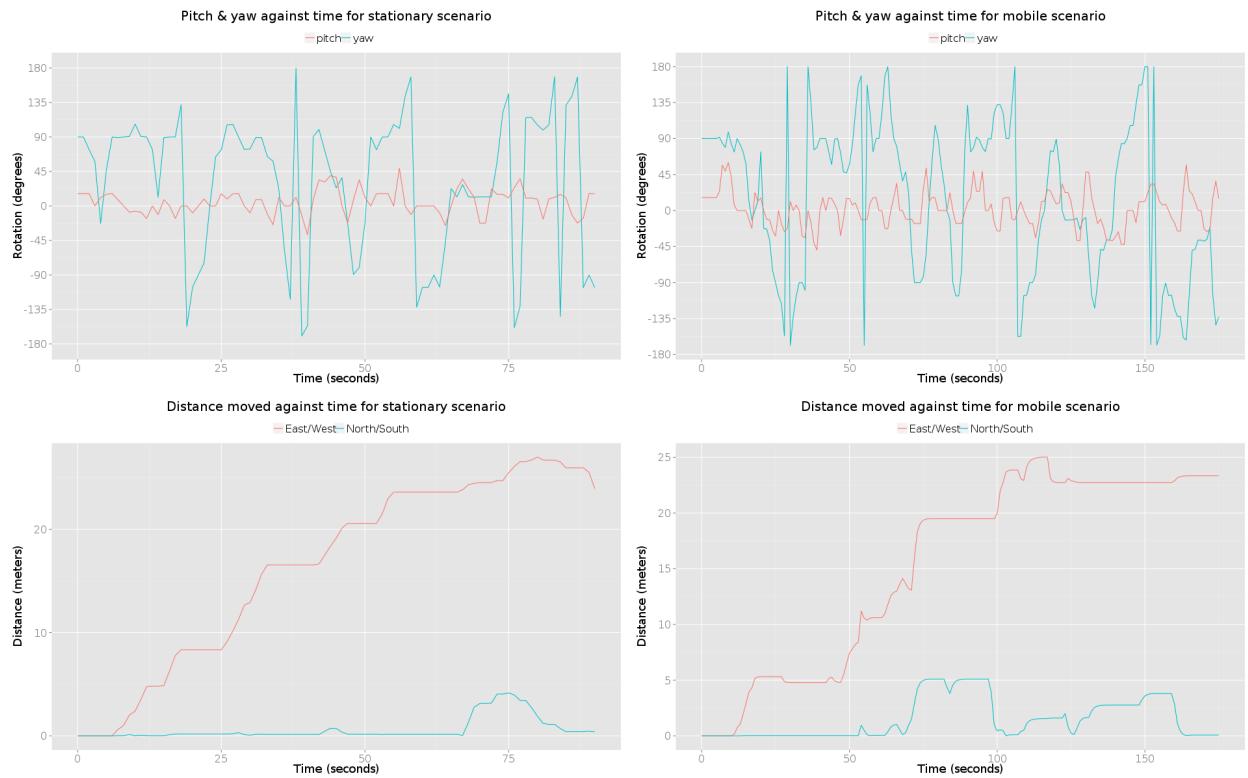


Figure 6.12: Participant 5

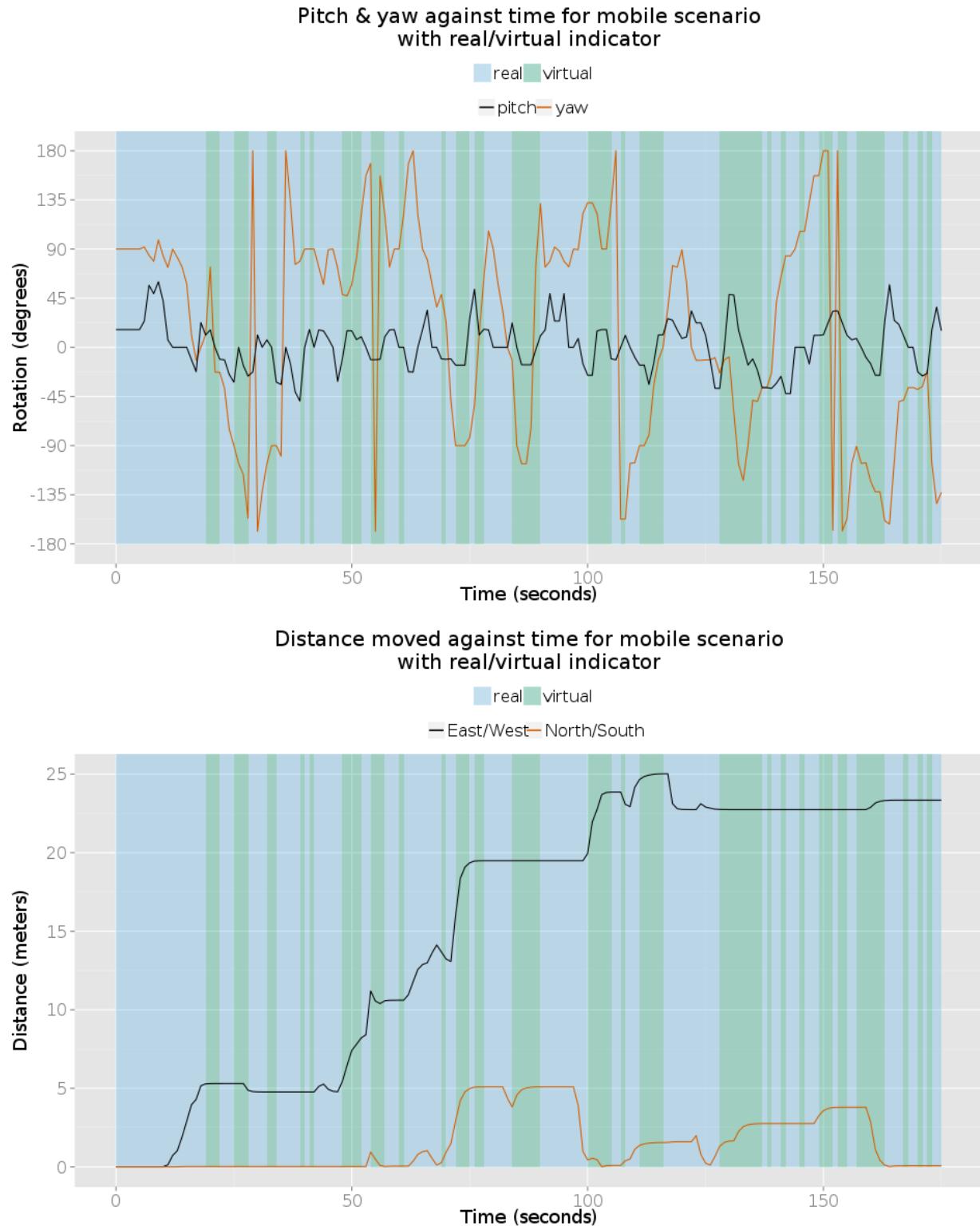


Figure 6.13: Participant 5



Pre Task Questionnaire

What is your age?

(leave blank if you prefer not to disclose)

How do you identify your gender?

(leave blank if you prefer not to disclose)

Do you have previous experience using a
games console controller?

Yes No

(circle applicable answer)

Do you have previous experience using a
Head Mounted Display (HMD)?

Yes No

(circle applicable answer)

Have you previously visited St Salvator's chapel?

Yes No

(circle applicable answer)

Have you previously interacted with the virtual
reconstruction of St Salvator's chapel?

Yes No

(circle applicable answer)

Researcher's use only

Date

Time



Post Task Questionnaire

University of
St Andrews

Please tick one box for each question

	Strongly disagree					Strongly agree				
I think that I would like to use this system frequently	<input type="checkbox"/>									
	1	2	3	4	5					
I found the system unnecessarily complex	<input type="checkbox"/>									
	1	2	3	4	5					
I thought the system was easy to use	<input type="checkbox"/>									
	1	2	3	4	5					
I think that I would need the support of a technical person to be able to use this system	<input type="checkbox"/>									
	1	2	3	4	5					
I found the various functions in this system were well integrated	<input type="checkbox"/>									
	1	2	3	4	5					
I thought there was too much inconsistency in this system	<input type="checkbox"/>									
	1	2	3	4	5					
I would imagine that most people would learn to use this system very quickly	<input type="checkbox"/>									
	1	2	3	4	5					
I found the system very cumbersome to use	<input type="checkbox"/>									
	1	2	3	4	5					
I felt very confident using the system	<input type="checkbox"/>									
	1	2	3	4	5					
I needed to learn a lot of things before I could get going with this system	<input type="checkbox"/>									
	1	2	3	4	5					

Researcher's use only

Date

Time



Post Task Questionnaire

University of
St Andrews

Please tick one box for each question

	Strongly disagree					Strongly agree					
	1	2	3	4	5		1	2	3	4	5
I found the exploration an enjoyable experience	<input type="checkbox"/>		<input type="checkbox"/>								
It was easy to compare features from the past & the present	<input type="checkbox"/>		<input type="checkbox"/>								
I felt motion sickness/dizziness	<input type="checkbox"/>		<input type="checkbox"/>								
In the virtual environment, I had a sense of being there	<input type="checkbox"/>		<input type="checkbox"/>								
I was aware of both real & virtual environments	<input type="checkbox"/>		<input type="checkbox"/>								
It was rewarding to explore the chapel in this way	<input type="checkbox"/>		<input type="checkbox"/>								
It felt as though I was in the past	<input type="checkbox"/>		<input type="checkbox"/>								
I think I would have preferred a conventional computer monitor	<input type="checkbox"/>		<input type="checkbox"/>								
This experience changed my understanding of the chapel	<input type="checkbox"/>		<input type="checkbox"/>								
I did not notice differences between the real & virtual environments	<input type="checkbox"/>		<input type="checkbox"/>								
The visual quality of the headset was bad	<input type="checkbox"/>		<input type="checkbox"/>								
I feel I now better understand what the chapel was like in the past	<input type="checkbox"/>		<input type="checkbox"/>								



Structured Interview

Encourage interviewees to elaborate on their answers to all questions (for example the response to question 1 should be more than just 'Scenario A' or 'Scenario B', question 4 should not be just 'yes' or 'no').

1. Which scenario did you prefer?
2. Which scenario did you find more engaging &/or rewarding?
3. Which scenario did you find made it easier to spot differences between the real environment & the virtual environment?
4. Did you notice any differences in scenario A that you did not notice in scenario B?
5. Did you notice any differences in scenario B that you did not notice in scenario A?
6. Did you experience any motion sickness/simulator sickness/dizziness whilst using the HMD? If so, was it better or worse in one scenario compared to the other?
7. Any further comments?

Conclusions & Discussion

- **Content** - More in-depth explanation of contributions, identify future work (but don't dwindle).
 - **What has been done** - Nothing.
 - **What is left to do/how long should it take** - Should probably be written towards the end, after all design & implementation, evaluation & discussion has been completed. Should be a short section, 10 pages is probably far too long, shouldn't take long to write if the evaluation & discussion section is done properly.
-

Appendices

Bibliography

- [1] Steven Warburton. Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology*, 40(3):414–426, 2009.
- [2] Beth Coleman. Using Sensor Inputs to Affect Virtual and Real Environments. *Pervasive Computing, IEEE*, 8(3):16–23, July 2009.
- [3] Raymond P Ten Eyck. Simulation in emergency medicine training. *Pediatric Emergency Care*, 27(4):342–344, 2011.
- [4] Hang Qiu and Leiting Chen. Real-Time Virtual Military Simulation System. In *2009 First International Conference on Information Science and Engineering*, number 2007, pages 1391–1394. Ieee, 2009.
- [5] Aaron Druck. When Will Virtual Reality Become a Reality?, 2006.
- [6] Yesha Sivan. 3D3C Real Virtual Worlds Defined : The Immense Potential of Merging 3D , Community , Creation , and Commerce 3D3C Real Virtual Worlds Defined : The Immense Potential of Merging 3D , Community , Creation , and Commerce. *Journal of Virtual Worlds Research*, 1(1):1–32, 2008.
- [7] Cisco. Cisco Visual Networking Index : Forecast and Methodology , 2010 – 2015. *Middle East*, June(1):2010–2015, 2011.
- [8] Faja Bose and Raja Bose. Sensor Networks Motes, Smart Spaces, and Beyond. *Pervasive Computing, IEEE*, 8(3):84–90, July 2009.
- [9] Paolo Baronti, Prashant Pillai, Vince Chook, Stefano Chessa, Alberto Gotta, and Fun Hu. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. *Comput. Commun.*, 30(7):1655–1695, 2007.
- [10] Paul Milgram and Fumio Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems*, E77-D(12):1321–1329, 1994.
- [11] Paul Milgram and Herman Colquhoun Jr. A Taxonomy of Real and Virtual World Display Integration, 1999.
- [12] Roy Want. Through Tinted Eyeglasses. *IEEE Pervasive Computing*, 8(3):2–4, July 2009.
- [13] Judy York and Parag C Pendharkar. Humancomputer interaction issues for mobile computing in a variable work context. *International Journal of Human-Computer Studies*, 60(5-6):771–797, 2004.
- [14] Joshua Lifton. *Dual Reality: An Emerging Medium*. Ph.d. dissertation, Massachusetts Institute of Technology, Department of Media Arts and Sciences, September 2007.
- [15] D W F Van Krevelen and R Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *The International Journal of Virtual Reality*, 9(2):1–20, 2010.

- [16] Farjana Eishita and Kevin Stanley. THEEMPA: simple AR games using layar. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, Futureplay '10, pages 219–222, New York, NY, USA, 2010. ACM.
- [17] María Caballero, Ting-Ray Chang, María Menéndez, and Valentina Occhialini. Behand: augmented virtuality gestural interaction for mobile phones. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, MobileHCI '10, pages 451–454, New York, NY, USA, 2010. ACM.
- [18] Joseph A. Paradiso and James A. Landay. Cross-Reality Environments. *Pervasive Computing, IEEE*, 8(3):14–15, July 2009.
- [19] Marie Kim, Hwang Jae Gak, and Cheol Sig Pyo. Practical RFID + sensor convergence toward context-aware X-reality. In *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, ICIS '09, pages 1049–1055, New York, NY, USA, November 2009. ACM.
- [20] Joshua Lifton and Joseph Paradiso. Dual Reality: Merging the Real and Virtual. In *Proceedings of the First International ICST Conference on Facets of Virtual Environments (FaVE)*, July 2009.
- [21] Responsive Environments Group. Responsive Environments Group - Research Themes.
- [22] J A Paradiso. The interactive balloon: Sensing, actuation and behavior in a common object. *IBM Systems Journal*, 35(3):473–487, 1996.
- [23] Phillip Rowe. *Characterization of a Wideband Monopulse Piezoelectric Direction Finder*. PhD thesis, Massachusetts Institute of Technology, 1999.
- [24] S E Burke and J A Paradiso. High-Resolution Piezopolymer Acoustic Bearing Estimator. In *Proc Second Technical Conference on Telecommunications Research and Development in Massachusetts electronic*, page 13, 1996.
- [25] Joseph A Paradiso and Eric Hu. Expressive Footwear for Computer-Augmented Dance Performance. In *First International Symposium on Wearable Computers ISWC 97*, pages 165–166. IEEE Computer Society, 1997.
- [26] Ara Knaian. *A Wireless Sensor Network for Smart Roadbeds and Intelligent Transportation Systems*. PhD thesis, Massachusetts Institute of Technology, 2000.
- [27] Z Teegarden. *Embedded Low-Power Wireless Sensor System: Design of a Software Radio Base Station*. PhD thesis, Massachusetts Institute of Technology, 2001.
- [28] Joseph A, Mark Feldmeier, and Joseph A Paradiso. Wireless Motion Sensors An Interactive Music Environment for Large Groups with Giveaway Wireless Motion Sensors. *Spring*, 31(1):50–67, 2007.
- [29] Hongshen Ma and J Paradiso. The FindIT Flashlight : Responsive Tagging Based on Optically Triggered Microprocessor Wakeup. *UbiComp 2002 Ubiquitous Computing*, pages 1–8, 2002.
- [30] Stacy J Morris Bamberg, Ari Y Benbasat, Donna Moxley Scarborough, David E Krebs, and Joseph A Paradiso. Gait analysis using a shoe-integrated wireless sensor system. *IEEE transactions on information technology in biomedicine a publication of the IEEE Engineering in Medicine and Biology Society*, 12(4):413–423, 2008.
- [31] Mathew Laibowitz and Joseph A Paradiso. Parasitic Mobility for Pervasive Sensor Networks. *Pervasive Computing*, 3468:255 – 278, 2005.
- [32] Jason LaPenta. *Real-time 3-d Localization using Radar and Passive Surface Acoustic Wave Transponders*. PhD thesis, Massachusetts Institute of Technology, 2007.

- [33] Joshua Lifton and Jay Lee. Media Matrix : Self-organizing Distributed Physical Database. *CHI 01 extended abstracts on Human factors in computing systems CHI 01*, pages 193–194, 2001.
- [34] J Paradiso, C Abler, K Hsiao, and M Reynolds. The Magic Carpet: Physical Sensing for Immersive Environments. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 277–278. ACM Press, 1997.
- [35] J A Paradiso, K Hsiao, J Strickon, J Lifton, and A Adler. Sensor systems for interactive surfaces. *IBM Systems Journal*, 39(3):892–914, 2000.
- [36] Bruce Richardson, Krispin Leydon, Mikael Fernstrom, and Joseph A. Paradiso. Z-Tiles. In *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04*, page 1529, New York, New York, USA, April 2004. ACM Press.
- [37] Joshua Lifton, Mark Feldmeier, Yasuhiro Ono, Cameron Lewis, and Joseph A. Paradiso. A platform for ubiquitous sensor deployment in occupational and domestic environments. In *Proceedings of the 6th international conference on Information processing in sensor networks - IPSN '07*, page 119, New York, New York, USA, April 2007. ACM Press.
- [38] Joshua Lifton, Manas Mittal, Michael Lapinski, and Joseph Paradiso. Tricorder: A mobile sensor network browser. In *Proceedings of the ACM CHI 2007 Conference - Mobile Spatial Interaction Workshop*, 2007.
- [39] Manas Mittal and Joseph A Paradiso. UbiCoder: A Mobile Device for Situated Interactions With Sensor Networks. *Ieee Sensors Journal*, 11(3):818–828, 2011.
- [40] Joshua Lifton. Consumer Adoption of Cross Reality Systems. In *Proceedings of the International Symposium on Ubiquitous Virtual Reality 2010 (ISUVR 2010)*, pages 9–11. IEEE, July 2010.
- [41] Ian Hughes. Lessons from second life, 2006.
- [42] Ian Hughes. Second Life Outside in, 2006.
- [43] Ian Hughes. Well it got my attention Second Life, 2006.
- [44] Ian Hughes. Wimbledon in Second Life, 2006.
- [45] Ian Hughes. Wimbledon A fanployee perspective, 2009.
- [46] Rita J King. From the Fire Pit. *New Writing*, 3(1):80–86, 2008.
- [47] Palmisano Gets A Second Life. Palmisano Gets A Second Life, 2006.
- [48] Paul Mason. Searching for the soul of cyberspace, 2007.
- [49] Gershon Dublon, Laurel S Pardue, Brian Mayton, Noah Swartz, Nicholas Joliat, Patrick Hurst, and Joseph A Paradiso. DoppelLab: Tools for exploring and harnessing multimodal sensor network data, 2011.
- [50] Gershon Dublon. *Beyond the Lens: Communicating Context through Sensing, Video and Visualization*. PhD thesis, Massachusetts Institute of Technology, 2011.
- [51] Joshua Lifton, Mathew Laibowitz, Drew Harry, Nan-wei Gong, Manas Mittal, and Joseph A. Paradiso. Metaphor and Manifestation - Cross-Reality with Ubiquitous Sensor/Actuator Networks. *IEEE Pervasive Computing: Mobile and Ubiquitous Systems*, 8(3):24–33, July 2009.
- [52] Nihal Dindar, Çagri Balkesen, Katina Kromwijk, and Nesime Tatbul. Event Processing Support for Cross-Reality Environments. *IEEE Pervasive Computing*, 8(3):34–41, July 2009.

- [53] Daniel Horn, Ewen Cheslack-Postava, Tahir Azim, Michael J. Freedman, and Philip Levis. Scaling Virtual Worlds with a Physical Metaphor. *IEEE Pervasive Computing*, 8(3):50–54, July 2009.
- [54] Justin Clark-Casey. *Scaling OpenSimulator: An Examination of Possible Architectures for an Internet-Scale Virtual Environment Network*. PhD thesis, Oxford, 2010.
- [55] Nicole Yankelovich, Jordan Slott, Alex Hill, Matt Bonner, Jacob Schiefer, Blair MacIntyre, Elena Mugellini, Omar Abou Khaled, Frédéric Barras, Jacques Bapst, Maribeth Back, Edgardo Aviles-Lopez, and J. Antonio Garcia-Macias. Building and Employing Cross-Reality. *IEEE Pervasive Computing*, 8(3):55–57, July 2009.
- [56] Jacques Fayolle. The Concept of the Remote Laboratory. \url{http://blogs.openwonderland.org/2011/01/28/remote-laboratories/}, January 2011.
- [57] UgoTrade. EOLUS Makes Leap To 3D Internet On Second Life. \url{http://www.ugotrade.com/2007/07/02/eolus-makes-leap-to-3d-internet-on-second-life/}, July 2007.
- [58] IBM. Made in IBM Labs: IBM 3-D Data Centers Show Virtual Worlds Fit for Business, 2008.
- [59] Marketwire. IBM 3-D Data Centers Show Virtual Worlds Fit for Business, 2008.
- [60] Linden Research Inc. Virtual Worlds, Avatars, free 3D chat, online meetings - Second Life Official Site.
- [61] OpenSimulator Project. OpenSim.
- [62] Open Wonderland Foundation. Open source 3D virtual collaboration toolkit — Open Wonderland.
- [63] Unigine Corp. Unigine: real-time 3D engine (game, simulation, visualization and VR).
- [64] Epic Games. Game Engine Technology by Unreal.
- [65] Unity Technologies. Unity - Game Engine.
- [66] Adel Hendaoui, Moez Limayem, and Craig W Thompson. 3D Social Virtual Worlds: Research Issues and Challenges. *IEEE Internet Computing*, 12(1):88–92, 2008.
- [67] William Sims Bainbridge. The scientific research potential of virtual worlds. *Science*, 317(5837):472–6, 2007.
- [68] Robert Faludi. *Building Wireless Sensor Networks*. O'Reilly Media Inc., Sebastopol, 1 edition, 2010.
- [69] TinyOS Alliance. TinyOS Home Page.
- [70] Adam Dunkels. The Contiki OS.
- [71] Mike Botts, George Percivall, Carl Reed, John Davidson, Silvia Nittel, Alexandros Labrinidis, and Anthony Stefanidis. GeoSensor Networks. *Lecture Notes in Computer Science*, 4540(175):175–190, 2008.
- [72] Inc. Open Geospatial Consortium. OWS-4 Demonstration.
- [73] International Telecommunication Union (ITU). ReqITU-T Y.2221 Requirements for support of ubiquitous sensor network (USN) applications and services in the NGN environment, January 2010.
- [74] Kyungeun Park, Jekuk Yun, Changhyun Byun, Yanggon Kim, and Juno Chang. XCREAM: collaborative middleware framework for RFID/USN-enabled applications. In *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services*, iiWAS '10, pages 692–695, New York, NY, USA, 2010. ACM.

- [75] Aman Kansal, Suman Nath, Jie Liu, and Feng Zhao. SenseWeb: An Infrastructure for Shared Sensing. *Multimedia, IEEE*, 14(4):8–13, October 2007.
- [76] Microsoft Research. SenseWeb Tutorial, 2008.
- [77] Karl Aberer, Manfred Hauswirth, and Ali Salehi. The Global Sensor Networks middleware for efficient and flexible deployment and interconnection of sensor networks. 2006.
- [78] Lee Kang and Kang Lee. IEEE 1451: A standard in support of smart transducer networking. In *Instrumentation and Measurement Technology Conference, 2000. IMTC 2000. Proceedings of the 17th IEEE*, volume 2, pages 525–528. IEEE, 2000.
- [79] Eugene Song and Kang Lee. Understanding IEEE 1451-Networked smart transducer interface standard - What is a smart transducer? *IEEE Instrumentation & Measurement Magazine*, 11(2):11–17, April 2008.
- [80] Wilfried Elmenreich, Stefan Pitzek, and Stefan Pitzek Wilfried Elmenreich. Smart Transducers - Principles, Communications, and Configuration. In *In Proceedings of the 7th IEEE International Conference on Intelligent Engineering Systems*, pages 510–515, 2003.
- [81] OMG. Smart Transducer Interface Specification, 2007.
- [82] Kang K.B. Lee and Richard R.D. Schneeman. Distributed measurement and control based on the IEEE 1451 smart transducer interface standards. *Instrumentation and Measurement, IEEE Transactions on*, 49(3):621–627, June 2000.
- [83] International Organization for Standardization and (ISO). ISO/IEC 23005-1:2011 - Information technology – Media context and control – Part 1: Architecture. \url{http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=54985}, 2011.
- [84] Yesha Sivan. Real Virtual Worlds SOS (State of Standards) Q3-2008. *Journal of Virtual Worlds Research*, 1(2), 2008.
- [85] Yesha Sivan. Overview: State of Virtual Worlds Standards in 2009. *Journal of Virtual Worlds Research*, 2(3), 2009.
- [86] Philip Rosendale. Virtual Worlds, Collaboratively Built. *Journal of Virtual Worlds Research*, 2(3), 2009.
- [87] Jean Gelissen. Introduction to MPEG-V. *Journal of Virtual Worlds Research*, 2(3), 2009.
- [88] Yesha Sivan. Managing Editors Corner: Celebrating Four Years and Planning for the Next. *Journal of Virtual Worlds Research*, 4(3), 2011.
- [89] Jean Gelissen, Marius Preda, Samuel Cruz-Lara, and Yesha Sivan. Issue Editors' Corner: The Current and Future Angles of Standards. *Journal of Virtual Worlds Research*, 4(3), 2011.
- [90] Jean Gelissen and Yesha Sivan. The Metaverse1 Case: Historical Review of Making One Virtual Worlds Standard (MPEG-V). *Journal of Virtual Worlds Research*, 4(3), 2011.
- [91] Jose Manuel Cabello, Jose Maria Franco, Antonio Collado, Jordi Janer, Samuel Cruz-Lara, David Oyarzun, Albert Armisen, and Roland Geraerts. Standards in Virtual Worlds Virtual Travel Use Case Metaverse1 Project. *Journal of Virtual Worlds Research*, 4(3), 2011.
- [92] IEEE. IEEE VW Standard Working Group.

- [93] Institute of Electrical and Electronics Engineers. IEEE P1828/D1 Draft Trial-Use Standard for Systems with Virtual Components. 2012.
- [94] C J Davies, Alan Miller, and Colin Allison. Virtual Time Windows : Applying Cross Reality to Cultural Heritage. In *PGNet*, 2012.
- [95] Wade Roush. Second Earth, 2007.
- [96] Avi Bar-Zeev. RealityPrime Second Earth, 2007.
- [97] David Gelernter. *Mirror Worlds: or the Day Software Puts the Universe in a Shoebox...How It Will Happen and What It Will Mean*. 1993.
- [98] Lars Qvortrup. Cyberspace as Representation of Space Experience: In Defence of a Phenomenological Approach. In Lars Qvortrup, editor, *Virtual Space: Spatiality in Virtual Inhabited 3D Worlds*, chapter 1, pages 5–24. Springer-Verlag, 2002.
- [99] E L Waterworth and J A Waterworth. 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, 2001.
- [100] S Laureys, M Boly, G Moonen, and P Maquet. Two Dimensions of Consciousness: Arousal and Awareness. 2:1133–1142, 2009.
- [101] Raffaele De Amicis, Gabrio Girardi, Michele Andreolli, Giuseppe Conti, and Raffaele De Amicis. Game based technology to enhance the learning of history and cultural heritage. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, ACE '09, page 451, New York, NY, USA, October 2009. ACM.
- [102] Steve Mann and James Fung. EyeTap Devices for Augmented, Deliberately Diminished, or Otherwise Altered Visual Perception of Rigid Planar Patches of Real-World Scenes, 2002.
- [103] Peter Bø gh Andersen. Tangible Objects : Connecting Informational and Physical space.
- [104] Kangsoo Kim, Byung-Kuk Seo, Jae-Hyek Han, and Jong-Il Park. 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*, page 323, New York, New York, USA, December 2009. ACM Press.
- [105] Rafal Wojciechowski, Krzysztof Walczak, Martin White, and Wojciech Cellary. Building Virtual and Augmented Reality museum exhibitions. In *Proceedings of the ninth international conference on 3D Web technology - Web3D '04*, page 135, New York, New York, USA, April 2004. ACM Press.
- [106] Fumio Okura, Masayuki Kanbara, and Naokazu Yokoya. Fly-through Heijo Palace Site : Augmented Telepresence Using Aerial Omnidirectional Videos. *Computer*, 1:4503–4503, 2006.
- [107] Nadia Magnenat-Thalmann, George Papagiannakis, and Parag Chaudhuri. Interactive virtual humans in mobile augmented reality. In *Encyclopedia of Multimedia*, pages 362–368. 2008.
- [108] Vassilios Vlahakis, Nikos Ioannidis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Luis Almeida, Didier Stricker, Tim Gleue, Ioannis T. Christou, and Renzo Carlucci. Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, VAST '01, pages 131–140, New York, NY, USA, November 2001. ACM.

- [109] Fumio Okura, Masayuki Kanbara, and Naokazu Yokoya. 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*, 2011.
- [110] Byung-Kuk Seo, Kangsoo Kim, Jungsik Park, and Jong-Il Park. A tracking framework for augmented reality tours on cultural heritage sites. In *Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry - VRCAI '10*, page 169, New York, New York, USA, December 2010. ACM Press.
- [111] Hrvoje Benko, Edward W Ishak, and Steven Feiner. Collaborative Mixed Reality Visualization of an Archaeological Excavation. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR '04*, pages 132–140, Washington, DC, USA, 2004. IEEE Computer Society.
- [112] George Papagiannakis, Sebastien Schertenleib, Michal Ponder, Marlène Arévalo-Poizat, Nadia Magnenat-Thalmann, and Daniel Thalmann. Real-time virtual humans in AR sites. In *1st European Conference on Visual Media Production CVMP*, number 4, pages 273–276. IET, 2004.
- [113] Nadia Magnenat-Thalmann, Alessandro E Foni, George Papagiannakis, and Nedjma Cadi-Yazli. Real Time Animation and Illumination in Ancient Roman Sites. *The International Journal of Virtual Reality2*, 6(1):11–24, 7.
- [114] George Papagiannakis, Sbastien Schertenleib, Brian O’Kennedy, Marlene Arevalo-Poizat, Nadia Magnenat-Thalmann, Andrew Stoddart, and Daniel Thalmann. Mixing virtual and real scenes in the site of ancient Pompeii. *Computer Animation And Virtual Worlds*, 16(1):11–24, 2005.
- [115] George Papagiannakis and Nadia Magnenat-Thalmann. Mobile Augmented Heritage: Enabling Human Life in Ancient Pompeii. *International Journal of Architectural Computing*, 5(2):396–415, 2007.
- [116] George Papagiannakis and Nadia Magnenat-Thalmann. Recreating Daily Life in Pompeii. *Arqueologica 2.0*, 1:16–20, 2009.
- [117] Takafumi Taketomi, Tomokazu Sato, and Naokazu Yokoya. AR Cultural Heritage Reconstruction Based on Feature Landmark Database Constructed by Using Omnidirectional Range Sensor. In *Computer Vision ACCV 2010 Workshops*, pages 265–275. 2011.
- [118] Krzysztof Walczak and Martin White. Cultural heritage applications of virtual reality. In *Proceedings of the eighth international conference on 3D Web technology, Web3D '03*, pages 182–183, New York, NY, USA, March 2003. ACM.
- [119] Maria Roussou, George Drettakis, and George Drettakis Maria Roussou. Photorealism and Non-Photorealism in Virtual Heritage Representation. In *Proceedings of the International Symposium on Virtual Reality, Archeology and Cultural Heritage (VAST)*, 2003.
- [120] Marcio Cabral, Marcelo Zuffo, Silvia Ghirotti, Olavo Belloc, Leonardo Nomura, Mario Nagamura, Fernanda Andrade, Regis Faria, and Leandro Ferraz. An experience using X3D for virtual cultural heritage. In *Proceedings of the twelfth international conference on 3D web technology, Web3D '07*, pages 161–164, New York, NY, USA, 2007. ACM.
- [121] Maria Roussou. Virtual Heritage : From the Research Lab to the Broad Public. In *VAST 2000 Euroconference*, pages 93–100. BAR International Series 1075, Oxford, Archaeopress, 2002.
- [122] Chris Christou, Cameron Angus, Celine Loscos, Andrea Dettori, and Maria Roussou. 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*, page 133, New York, New York, USA, November 2006. ACM Press.

- [123] J. Willmott, L. I. Wright, D. B. Arnold, and A. M. Day. 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*, VAST '01, pages 111–120, New York, NY, USA, November 2001. ACM.
- [124] Delia Tzortzaki. Virtual Reality as Simulation: The CAVE as "Space of Illusion" in Museum Displays. In Lars Qvortrup, editor, *Virtual Space: Spatiality in Virtual Inhabited 3D Worlds*, pages 258–284. Springer-Verlag, 2002.
- [125] IndoorAtlas Ltd. Ambient magnetic field-based indoor location technology. 2012.
- [126] John Brooke. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189:194, 1996.
- [127] Mel Slater. Presence and The Sixth Sense. *Presence: Teleoperators and Virtual Environments*, 11(4):435–439, August 2002.
- [128] Wijnand Ijsselsteijn. Understanding Presence. In *Proceedings of the AIIA 2002 'Workshop sulla percezione della presenza in ambienti virtuali o remoti'*, 2001.
- [129] Carrie Heeter. Reflections on Real Presence by a Virtual Person, 2003.
- [130] Wijnand Ijsselsteijn, Huib de Ridder, Roelof Hamberg, Don Bouwhuis, and Jonathan Freeman. Perceived depth and the feeling of presence in 3DTV, 1998.
- [131] Sherry Turkle. Our Split Screens. In Andrew Feenberg and Darin Barney, editors, *Community in the Digital Age: Philosophy and Practice*, chapter 6, pages 101–117. Rowman & Littlefield, 2004.
- [132] Mel Slater and Anthony Steed. A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments*, 9(5):413–434, October 2000.
- [133] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The Experience of Presence: Factor Analytic Insights, 2001.