PRESENCE, IMMERSION AND THE PANORAMA

A theoretical, technical and artistic inquiry into the nature of presence and immersion in virtual reality.

BY

MATTHEW M^C^GINITY

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy

THE UNIVERSITY OF NEW SOUTH WALES

SCIENTIA

MANU ET MENTE

SYDNEY · AUSTRALIA

School of Media Arts, College of Fine Arts

March 2014

ABSTRACT

In this thesis, the phenomena of immersion and presence are explored through three bodies of work: a theoretical investigation, the construction of an immersive panoramic virtual reality theatre, and the creation of an interactive immersive artwork.

In the first part, the concept of *immersive mimesis* is introduced to understand the assumptions and ramifications of the very idea that the "being there" we enjoy in the real world might be possible in a mediated experience. Drawing on J. J. Gibson's ecological approach to perception, presence is identified as *active perception of a light field*, a notion which is further refined to the *act of creating and detecting invariant structures in multi-modal stimuli.* This framing of presence serves as a common basis for understanding the immersive roles of a variety of perceptual phenomena, including the 10 degrees of freedom of vision, ecological optics, stereoscopy, ego-motion, vection and perceptual rest-frames, binding of stimuli into singular percepts, cross-sensory enhancements and transfer, interaction and perception of causality, and the destructive effect of the image when used as a surrogate for the light field.

The second part involves the creation of a panoramic multi-user immersive theatre based not only on contemporary virtual reality techniques and technologies, but on the understanding of immersion and presence arrived at above. Four pivotal features distinguish it from its panoramic heritage: omnistereoscopic imagery, spatial audio, real-time computation and interactivity.

Finally, in order to explore the immersive and aesthetic potential of this new incarnation of the panorama, a work of art is conceived. *La Dispersion du Fils*, an algorithmic invocation of the tragedy of Actaeon, takes the form of a never-ending, never-repeating voyage through fields and structures constructed wholly from moving images and sound. In this work, all the elements of the theory of presence developed above are explored, demonstrating that contemporary presence theory can inform artistic creation and that the pursuit of an art of immersion can provide insight into the nature of presence.

Originality statement

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed ....

Date ....18..09..2014...............................

Copyright statement

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed ....

Date ....18..09..2014.........................

Authenticity statement

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed ....

Date ....18..09..2014........................

ACKNOWLEDGEMENTS

I must begin by expressing my deepest gratitude to my thesis supervisor, Professor Ross Harley. Whenever doubt bettered me, Ross would fill my sails and set me back on course. Without Ross's unflagging optimism, enthusiasm, kindness and support, this work would never have reached completion.

I am also indebted to my former colleagues at the iCinema Centre: Joachim Tesch, Keir Smith, Jared Berghold, Alex Kupstov, Balint Seeber, Ardrian Hardjono, Robin Chow, Som Guan, Marc Chee, Rob Lawther, Anuraag Sridhar, Volker Kuchelmeister, Sue Midgley, Damian Leonard, Kate Dennis, Densan Obst, Joann Bowers and Professors Dennis Del Favero, Maurice Pagnucco and Paul Compton. The iCinema projects mentioned within these pages were born from the united talents of all these remarkable people. Together, it seems, we have achieved a great deal.

I would like to thank Professor Jeffrey Shaw, the visionary force behind the AVIE system and the iCinema Centre as a whole, without whom none of the work in this thesis would ever have taken place. It was Jeffrey who introduced me to the world of immersive art, and it is impossible to overstate the influence he has had on my understanding of the genre.

Sincere thanks must also go to Richard Castelli, director of *Epidemic Productions*, for initiating, funding, producing and organising the numerous exhibitions of *La Dispersion Du Fils*. Without Richard, *La Dispersion du Fils* would not exist.

I consider myself truly fortunate to have had the opportunity to work with Jean Michel Bruyère, Delphine Varas and Thierry Arredondo. Working with these extraordinary artists revealed to me new worlds, new ways of working, and the true nature of artistic inquiry. I consider our deep and continuing friendship to be the most valuable outcome of this thesis.

To my family and friends, I give thanks for their infinite patience.

Finally, I dedicate this work to my dear Sladja, whose sacrifice shall never be forgotten, and to little Mila, who was born somewhere around page 100.

This research was supported by an Australian Research Council Post Graduate Award.

Supplementary Material

Video documentation, additional images and an electronic version of this text can be found online at *www.ladispersiondufils.net.*

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

1. INTRODUCTION

In 1787, Robert Barker wrote in the patent awarded for his invention *La Nature à Coup d'Œil* that the intended purpose of his apparatus was "to make observers, on whatever situation he may wish they should imagine themselves, feel as if really on the very Spot." The apparatus in question was the panorama,[1] or cyclorama as it would later be known, and it was designed specifically to provoke in the audience a sensation of "being there," or, as we might say in contemporary parlance, a state of immersion or a sense of presence.[2] This desire to instil in the audience the sensation of "being there" can be traced throughout the history of panoramic displays. From the numerous battle-themed cycloramas of the 19th century, through the extraordinary experiments in panoramic cinema at the dawn of the 20th century, the amusement parks and world fairs of 1960s, to the most recent digital incarnations, the panoramic form has repeatedly been employed in the pursuit of immersion. That the panorama is somehow associated with immersion, or a heightened sense of "being there" is, it would seem, both obvious and indubitable.

But what exactly is the relationship between the panoramic form and immersion?

This question, in various ways, represents the kernel of this thesis. In the manner of a sprouting seed, this question yields numerous lines of inquiry; lines that traverse rich and disparate fields of knowledge, from aesthetics to computer science, engineering and cognitive science. First and foremost, in asking what is immersive about the panorama, one must immediately confront the question of what is meant by the terms presence and immersion, a question that pervades the entirety of this thesis. This question, taken alone and without context, rapidly grows overwhelming, and so the concept of the panorama serves here as a framing device, a point of origin and return to

---

[1] Throughout the entirety of this text, the word 'panorama' is used strictly in reference to cylindrical images viewed from within, and should not be understood to mean wide-format or wide-field of view images in general, or any of its more figurative meanings.

[2] Precise definitions for the terms *immersion* and *presence* and the distinction between them are given in Section 2.2.2 below. Until then, the terms will be used interchangeably to denote a general sense of "being there."

be always kept in sight, lest we get lost among the innumerable paths of inquiry on offer.

Within this thesis, the relationship between the panorama and presence is explored through three bodies of work. The first is a theoretical investigation into the nature of presence and immersion. The second is the creation of the *Advanced Visualisation and Interaction Environment* or *AVIE*, a modern-day reimagining of the panorama. The third is *La Dispersion Du Fils,* an artwork created specifically to explore the immersive and aesthetic potential of the AVIE system. Together, these three components allow for an inquiry into the immersive potential of the panorama, and the nature of immersion and presence in general, from three perspectives: theoretical, technical and artistic.

The theoretical aspect of this work begins with the concept of *mimetic immersion*, a framework that serves to elucidate the basic assumptions underlying the very idea that presence might be attained in a mediated world. In doing this, presence is identified as an everyday feature of our being in the real-world, rather than something unique to mediated experiences. The presence we enjoy in the real-world is, therefore, the standard by which we measure presence in a mediated environment; and before aiming to replicate it in an artificial world, it must first be understood in the real-world.

It will be argued that a definition of presence resides in the identification of the aspects of our being in the real-world that must be replicated in the virtual world, if presence is to be enjoyed there too. That there might exist certain aspects-of-being necessary or sufficient for presence for all possible virtual worlds and viewers can be seen as a form of *presence hypothesis*, and the ramifications of and evidence for this hypothesis are considered. This touches on the fundamental connection between immersion and simulation and the question of what it might mean to simulate a fictional or non-existent world.

In this thesis, presence will be identified as a particular form of perceptive relationship with the environment. To understand this relationship, the reciprocal concepts of the *light field* and the *plenoptic function* are introduced. As all acts of vision and image capture, rendering, and display can be described in terms of the plenoptic function, it serves as a powerful tool for the analysis of visual media and our relationship with it. Here, it allows us to see that normal vision is an act of detecting change and structure in *all* dimensions of the plenoptic function, and that much of our ability to make sense of the world is dependent on continuous *ego-motion* – a seeing the world not from a point, but from a path. This will be shown to be the fundamental basis of perception-as-action, and it is this perceptive relationship with the world that is taken here as the basis of presence.

This picture of vision shares much in common with the ecological approach to vision of J. J. Gibson, and it is from Gibson that the notions of ecological optics and invariants are drawn. Vision, according to Gibson, is "seeing the non-change underlying the change" Gibson (1971, p. 32), an observation that reveals the true importance of movement in perception: only with variance can there be invariants. To this, the phenomena of *binding* – the fusion of disparate stimuli into single, unified percepts – is added, arriving at a notion of active perception as *the act of creating and detecting invariant structures in multi-modal stimuli.*

Armed with this definition of presence, the question of immersive displays is then addressed. A display can be considered immersive if it permits this same kind of active-perceptive relationship with a mediated light field - a definition that will be shown to yield a number of results. A display must permit exploration of the light field in all dimensions, preserve invariant structure, permit binding of stimuli (within and across sensory modalities), and allow perception of the virtual world as a *perceptual rest-frame*; all concepts that will be expounded in detail. The degree to which a display satisfies these criteria provides a measure of its immersiveness, and it will be shown that above all it is the *degrees of freedom* with which the viewer is free to navigate the plenoptic function that most influences sensations of presence.

At this point discussion turns to the perception of images. What kind of visual relationship with the virtual world is afforded by the image? What elements of vision are lost or distorted when an image is adopted as a surrogate for a light field? It will be shown that the vision of the world offered by an image is in fact nothing like that ordinarily enjoyed in the real world. For, in collapsing the light field to an image, all the degrees of freedom with which we might normally explore our visual environment are extinguished. Perception ceases to be an act, the body is excised from the equation and presence is lost. Further, in offering a dual-vision - one of the virtual world and one of the image-surface itself - the image introduces false invariants while destroying true ones. It is the image that presents the single greatest obstacle in the pursuit of an immersive display.

How then, given the pernicious effect of the image on immersion, might an image be used to construct an immersive display? Can any of the artefacts and distortions introduced by the image be diminished? Can any of the degrees of freedom of vision be somehow reinstated in the act of perception?

Further, what are the consequences of sharing an image between multiple viewers? How might a single image be used to immerse multiple viewers in the same shared virtual world? Is not the very idea of a shared immersive image a paradox?

It is in answering these questions that discussion turns to a variety of perceptual phenomena including stereoscopy and multiple-centre of projection images, cross-sensory transfer and enhancement, vection and perceptual rest-frames. Stereo images, combined with high resolutions and fields of view are shown to reduce the visibility of the image surface, a major source of false visual cues. More exotic image projections, where the point-of-view is not a point but a path, can be used to convey valid stereoscopic views of a world to multiple viewers in a single shared image. Multisensory 'synaesthetic' phenomena, where stimuli from one sensory channel enhance, alter or induce perception of stimuli in another, can be adopted as effective tools for reducing the visibility of flaws in the mediation. Simulated ego-motion – the simulation of a moving point of view – will be shown to offer some of the benefits of real motion, mitigating the loss of the spatial degrees of freedom of vision and reducing the perception of image distortion. In particular, vection – the perception of self-motion induced by visual stimuli – is shown to be intimately related to presence, a connection made explicit through the concept of a perceptual rest-frame.

In addition, the importance of interaction in the invocation of presence is considered. The immersive potential of interaction is explained in terms of causal perception – the perception of cause and effect – and interaction is shown to be an avenue to presence only so far as it gives rise to the perception of causality and active-perception of the environment. An important result that follows is the potentially destructive effect that interaction may have on presence, for while interaction can lead to greater immersion, *it can just as readily destroy it*. This undermines the generally unquestioned view that interactive experiences are necessarily more immersive than their passive counterparts.

These are some of the results arising from this theoretical inquiry into immersion and presence. Framing presence as active-perception of invariant structure in multi-sensory stimuli, the immersive roles of a surprisingly wide variety of perceptual phenomena are accounted for and united in a single explicative framework. To demonstrate the effectiveness of this framework, and to show that many of these ideas can in fact be found tacitly incorporated in artistic understanding of presence, the immersive creations of Cardiff and Miller, CREW and Jeffrey Shaw are briefly examined, bringing the first component of the thesis to an end.

The second body of work concerns the *Advanced Visualisation and Interaction Environment* or *AVIE*. Barker's original panorama was intended as nothing less than a device for virtual reality, conceived in terms of 18th century technology. The AVIE is a 21st century re-imagining of the panorama based not only on contemporary virtual reality techniques and technologies, but on current understanding of presence. Like its forebears, it is designed to immerse multiple viewers in a single shared space.

What, if anything, is immersive about the panorama? What perceptual mechanisms are at work? What might a modern incarnation of the panorama look like? What new immersive faculties are made possible with modern technology? Is the panoramic device effective in a 21st century pursuit of immersion, or has it long since become an atavism?

The design, development and subsequent deployment of the AVIE provides opportunity to answer such questions and to put the proposed theory of presence into practice. The result is a cylindrical immersive display, with four features that distinguish it from its panoramic ancestors: omnistereo imagery, spatial audio, real-time image generation and interactivity.

Images displayed in the AVIE are omnistereoscopic; a specific form of cylindrical projection that yields an effectively undistorted stereo image in all viewing directions, providing multiple viewers with the freedom to turn and explore the encircling environment at will. The omnistereo projection is strictly only correct for a viewer standing at the centre of the AVIE, and as the viewer moves about, the world appears to compress, dilate and shear. However for a stationary observer, curiously, distortions are largely imperceptible no matter where the viewer stands. The 'robustness of perspective,' a perceptual phenomena that accounts for the undistorted appearance of 2D perspective images no matter where they are viewed from, is discussed in attempt to understand this intriguing feature of the AVIE.

While the early panoramas often had sound effects or musical accompaniment, the spatial surround audio system of the AVIE permits accurate placement and movement of sounds within and around the space. It does this by exploiting a perceptual phenomenon known as *summing localisation*, a form of perceptual binding where two or more sounds arriving from different directions can, under certain specific conditions, be heard as a single 'phantom' sound emanating from a single position in space. In this sense, it is an example of what is termed here *metameric mimesis*, an approach to immersion that employs perceptual metamers – dissimilar physical stimuli that nevertheless give rise to indistinguishable perceptions – to reduce the complexity of the simulation.

The spatial sound system permits the creation of virtual worlds rich with spatially coherent visual and sonic cues, inviting such multi-sensory interactions as binding and cross-modal enhancement and transfer, all previously identified as playing roles in the generation of presence.

The AVIE is equipped with a vision based tracking system, allowing the design of interaction strategies based on the movement of viewers within the AVIE. This provides an ideal platform to study the interplay between interaction on immersion, and understand the circumstances where interaction serves as an aid to immersion, and those where it is a hindrance.

The fourth feature that distinguishes the AVIE from its panoramic heritage arises from its capacity to display images generated in real-time. This permits a completely different inquiry into immersion: what are the roles of procedural or generative algorithms and stochastic processes in the construction of immersive experiences? Real-time computation transforms the panorama from cinema into simulator; a machine in which worlds or narratives are no longer revealed or replayed, but constructed or discovered; a vessel for the exploration of landscapes and structures crafted not by hand, but by machines.

Presented here is a detailed technical account of the various components of the AVIE: the projection system used to illuminate a cylindrical screen with a single seamless and undistorted cylindrical image, the spatial surround sound system, the small cluster of computers synchronised to work as one, the input devices, and the iCinema *Software Development Kit (SDK)*, a platform for accelerating the development of AVIE applications.

The third and final body of work in this thesis is *La Dispersion Du Fils*, a work of art designed specifically for display in the AVIE. Created in collaboration with Jean Michel Bruyère, Delphine Varas and Thierry Arredondo, members of the Marseille-based artist collective LFKs, *La Dispersion Du Fils* is the continuation and climax of their decade long inquiry into the myth of Actaeon and Diana. *La Dispersion Du Fils* takes the form of a never-ending, never-repeating voyage through landscapes and structures constructed entirely from moving images, all drawn from the vast cinematic library of LFKs. The work is generative and stochastic, and places great emphasis on the roles of chance and discovery.

With its themes of metamorphosis, transcendence, the pursuit of knowledge and paths to impossible destinations, the work is designed to explore the artistic possibilities and limits of this modern panoramic form, and in doing so, validate or invalidate the very concept of the AVIE and the assumptions it embodies.

In creating a work of art, the theories and observations regarding presence and immersion described above are opened to an artistic perspective. For while presence research is by nature interdisciplinary, the language of presence research remains predominately scientific; it is the language of psychology, psychophysics and cognitive science, computer graphics and human-computer interaction, engineering and communications. Conspicuously missing from presence research is the language of art. *La Dispersion Du Fils*, as an artistic endeavour, allows inquiry into the aesthetics of presence. Through it, the aesthetic possibilities of concepts such as ecological optics, vection and ego-motion, perceptual rest-frames, omnistereo, spatial audio and synaesthesia, interaction and the perception of causality are explored. And in doing so, a bridge is built between artistic understanding of these phenomena and modern presence research.

The work also opens new avenues of inquiry, such as the re-purposing of traditional video material in an immersive environment, or the use of ego-motion to effect spatiotemporal montage and sonic composition, or the use of immersion as a 'fitness function' for algorithmic exploration of stochastic or generative systems, or a form of virtual surrealism where the viewer's trajectory through the space of all states traces a path on the thresholds of immersion or plausibility. Such aspects of the work can be considered the first signs of an emerging aesthetics of immersion.

1.1 APPROACH AND METHODOLOGY

The theoretical component of this thesis arises largely in response to the practical experience of constructing the AVIE, *La Dispersion Du Fils*, and a number of other immersive artworks developed at the iCinema Centre over a period of ten years. It is an attempt to explain presence and immersion as they have been observed, in practice. It is also an attempt to explain presence in immersion from first principles, beginning with a theory of how we see the world, and extending this to virtual environments. In identifying presence as a purely perceptual phenomena, it draws foremost on results arising from the study of vision and, to a lesser degree, hearing.

The technical component concerns the construction of an immersive and interactive display, and concerns the fields of stereoscopic projection, real-time computer graphics, spatial audio, computer clusters and distributed software. It does not set out to necessarily develop or further techniques in any of these sub-domains, but to create something entirely new by synthesising existing techniques. However, considerable effort is taken to provide introductions to each of these constituent fields of research, to both justify the design decisions made and in order that this text may serve as a form of AVIE 'user's manual' for future developers.

Practice-based Research

> There are circumstances where the best or only way to shed light on a proposition, a principle, a material, a process or a function is to attempt to construct something, or to enact something, calculated to explore, embody or test it. (Archer, 1995)

This thesis is 'practice-based.' Knowledge about presence is gained through the realisation and exhibition of a work of art. The lines of inquiry presented in the introduction are pursued through the process of creating an artwork - a process entailing experimentation and trial-and-error, construction of prototypes, exhibition and public scrutiny, and perhaps most importantly, chance and discovery.

The artwork developed within this thesis, *La Dispersion Du Fils,* serves foremost as a laboratory for research and experimentation. As such, it resides in a persistent state of flux, being continually reshaped to address different ideas, concepts and questions. Or rather, new layers are continually added upon old; each version a palimpsest record of earlier experimentation. As a work of art, it has no preconceived terminus.

As the knowledge gained from this research is encapsulated in the artwork itself, and communication of knowledge is the basis of research, public exhibition of the work is a vital part of this thesis. To this end, *La Dispersion Du Fils* has been exhibited on ten separate occasions.[3] It has never been exhibited in the same state twice, but always in a state reflecting the current progress "in the lab."

Beyond communication of results, there are further benefits arising from public exhibition of the work. For one, public exhibition demands that the work meet high standards of robustness and quality, and that adequate attention is given to both detail and the work as a whole. When exhibited publically, the work itself must tell the whole story; stand alone, so to speak. It demands that the experiments in immersion be contextualised thematically and housed in a narrative (in the loosest sense of the term).

This contextualisation is often in stark contrast to the manner in which immersion and presence are investigated in the laboratory, where very often immersive phenomena are studied by reducing scenario to the minimal set of stimuli required to trigger some perceptive response. Certainly this approach has been fruitful. However, it is hard not to see parallels between the rarefied and highly controlled laboratory tests that at one time dominated the field of cognitive psychology and current methods in presence research, in which case the same criticisms of lack of 'representative design' or 'ecological validity' levelled at the field of psychology by the likes of Brunswik (1956), Gibson (1966, 1979) or Neisser (1976) could equally be made of contemporary presence research. This is especially the case when a definition of presence predicated on Gibson's theory of ecological perception is adopted, as is the case in this thesis.

It is suggested that the artwork presented in this thesis, and many other immersive artworks alongside it, can be taken as psychophysical experiments into the nature of presence. And if they are, then they are experiments in which the subject is immersed in worlds that are far richer in sensory structure and complexity, and somehow more complete, than anything that might be found within traditional presence research.

Practice-led Research

This thesis is also, to a certain extent, "practice-led."[4] It is an examination of the practice of creating an immersive experience. The knowledge arising from this

---

[3] See Appendix A for details of exhibitions.

examination, encapsulated here in text and in the artwork itself, is immediately applicable to the creation of new works of art for the AVIE and for other immersive systems, panoramas or not.

Collaboration and Interdisciplinarity

Both the AVIE and the work *La Dispersion Du Fils* are the results of collaboration. The AVIE was developed by a small team of software and hardware engineers over many years of research at the iCinema Centre. As the chief computer scientist at iCinema throughout this period, and the lead developer of the *iCinema Software Development Kit (SDK)* - the software that drives it - the author played a major role in the creation of the AVIE. Approximately 80% of the code comprising the iCinema SDK was written by the author himself, and the parts that were not were written under the author's direct supervision. As such, this thesis largely concerns contributions to the AVIE system and iCinema SDK made by the author himself, but when it does not, care is taken to acknowledge authorship. In particular, the chapter concerning the AVIE contains analyses of components that were designed or implemented with or entirely by colleagues. In these cases, it is the analyses themselves which are presented as unique work, all of which are solely the work of the author.

La Dispersion Du Fils is the result of a collaboration with the three artists Jean Michel Bruyère, Delphine Varas and Thierry Arredondo, members of the Marseille-based artist collective LFKs. Here, each member played distinct roles in the creation of the artwork, and this too is duly noted throughout the thesis.

The collaborative nature of this work directly influences the outcomes of this research. First, it would not be possible for a researcher to cover so much ground alone. To create the AVIE, develop the software platform and then use it to construct an artwork such as *La Dispersion Du Fils*, all without collaboration, might represent a lifetime of work. Secondly, this thesis is predicated on an interdisciplinary study of immersion, and collaboration is the most effective way of pursuing such research. *La*

[4] I have adopted the terminology provided by Candy (2006): "If a creative artefact is the basis of the contribution to knowledge, the research is practice-based. If the research leads primarily to new understandings about practice, it is practice-led."

Dispersion Du Fils is the product of collaboration between computer scientists and engineers, film-makers, sculptors, writers and musicians.

This approach, however, is not without its problems. For one, there is the problem of language and communication of results. A research discipline provides not only a common language, but also a common knowledge, and an understanding of what needs to be explicitly stated and what does not. In resisting classification and attempting to straddle the three domains of computer science, engineering and art, this problem of language is repeatedly encountered throughout this text. For example, what language should be used to describe the form of the *Helix* - the principal character in *La Dispersion Du Fils*? Should it be described according to the algorithms employed, or in terms of aesthetics? And if we attempt both, what starting point can be assumed as common knowledge?

The approach adopted here is to resist classification. If inquiry into immersion must necessarily be interdisciplinary, then an interdisciplinary language and understanding will need to be developed. This challenge can be considered yet another path of inquiry open to this thesis, and this text one possible approach to the problem.

But, in the end, the approach adopted here will be a pragmatic one: I will write for myself and people of my ilk, and adopt the language of a computer scientist who works in the arts. This is, after all, the only language at my disposal.

Software Development

Behind this manuscript, and behind the artwork *La Dispersion Du Fils*, lies an enormous amount of custom software. It is the bricks and mortar of this thesis, and the development of this software accounts for majority of time and energy expended. In this respect, this research can be considered an exercise in software development. This work, however, resembles an iceberg, in that nine-tenths of it is hidden from view. For example, a component of this thesis is the use of traditional two-dimensional film to create three-dimensional space. While the aesthetic potential of this concept will be discussed in so far as it pertains to the research themes of the thesis, relatively little attention will be given to the creation of the system that made it possible, despite the hundreds of hours of software development this system required.

This touches on another challenge, and opportunity, inherent in an interdisciplinary inquiry into immersion. The vast majority of technical work required in the creation of the AVIE and its software revolves around integration, (almost always with modification), of existing technologies and techniques. This form of work, however, can be difficult to publish through traditional computer science channels. Despite the success of the AVIE platform, if measured by the projects it has given rise to and the construction of twelve AVIE replicas at different sites around the world, the technical work behind it has yielded extremely few publications through traditional computer science forums to date. Rather, the AVIE has spawned numerous publications in the fields of media art, and enjoyed frequent public exhibition. In short, projects such as the AVIE and *La Dispersion Du Fils*, which in practice are predominately exercises in software engineering, would simply not be possible without the artistic framework made viable by an interdisciplinary approach to immersion.

What this thesis is not

1. The thesis does not concern the history of the panorama, its social relevance or its influence on the development of cinema, art or virtual reality. These topics have been thoroughly explored elsewhere, notably in the work of Erkki Huhtamo and Oliver Grau.[5]

2. With regards to the artwork, this thesis focuses on the process of creation, rather than the result itself. It therefore does not attempt to methodically measure the 'success' of the artwork by evaluating visitor feedback, or by any other means.

Development of ideas

Finally, it is important to remember that this thesis tells a story. It is a retrospective account of the journey begun with the conception of the AVIE and ending with full realisation of *La Dispersion Du Fils*, and the many chance encounters and accidental

[5] See Huhtamo (2013), Grau (1999, 2003, 2004, 2007) and Grau and Veigl (2011).

discoveries in between. Some of the ideas presented in this text preceded the *La Dispersion Du Fils*, or were developed in concert, while others have been divined in retrospect, with the gift of hindsight.

1.2 Contemporary Relevance

> As technological developments continue to bring immersive technologies within the grasp (both economically and practically) of a growing user base, the need for understanding the properties and limitations of immersive experiences will increase proportionally. (Fisher et al., 2005, p. 632)

While the panorama, as a form of entertainment medium, enjoyed a golden age of popularity in the 18th and 19th centuries, it certainly does not today. The panorama survives only in small numbers, as research platforms or experimental prototypes, or curiosities; they are no longer considered viable platforms for mass entertainment.[6] Immersion, however, is witnessing a boom. For it could be argued that it is the unspoken desire for ever greater immersion that is the unwavering force behind the unrelenting evolution of media technology. What else might explain the seemingly insatiable market for televisions with higher resolution, cameras with more pixels, games consoles with more polygons, movies with more convincing artifice, if not some draw towards ever higher states of immersion? Whatever the cause for this pervasive (and largely unquestioned) compulsion for more immersive experiences, as the technology for immersive experiences advances and further permeates our lives, a greater and more nuanced understanding of immersion will be required.

1.3 Thesis Contributions

Within this thesis the reader will find three bodies of work, each of which contain novel contributions to the understanding or creation of immersive mediated experiences.

[6] The tradition of grand-scale panoramic painting is however, very much alive in the work of Yadegar Asisi, who in 2012 alone exhibited not one, but two enormous panoramic paintings in Berlin. See *www.asisi.de.*

In the first body of work, a novel theoretical account of presence is offered. This begins with the introduction of *mimetic immersion*, a framework within which the underlying assumptions behind the concepts of presence and immersion are made explicit, including the declaration of the *presence hypothesis* - the very idea that a sense of 'being there' in an *imperfectly* mediated world might be possible. This framework provides a concrete grounding for subsequent reasoning about presence and immersion, and to the author's knowledge is quite unique in presence literature.

Following this, a definition of presence is offered, in which J. J. Gibson's approach to ecological perception is combined with such concepts as multi-sensory binding, perceptual rest-frames and active-perception of the plenoptic function. While sharing some features with other existing definitions and theories, this synthesis of ideas is certainly unique, and (from the author's point of view) offers a more satisfactory and complete account of spatial perceptive presence than can be found in current presence literature. The manner in which this proposed model of presence accommodates and connects perceptive phenomena of ego-motion, vection and perceptual rest-frames, interaction, perception of causality, cross-modal interactions, degrees of freedom of vision, perception of images, spatial updating, is, to the author's knowledge, original.

The second major contribution of this thesis is the AVIE system. The AVIE represents a novel solution to the problem of immersing multiple viewers in a shared virtual space, and at the time the system was implemented, was perhaps the only example of a real-time immersive VR theatre designed expressly for large numbers of viewers. The combination of 360° omnistereo projection, spatial surround sound, real-time image generation and motion-based interaction, represents a novel synthesis of technologies, which until the AVIE, seem to have never featured together in a single system. Presented here, for the first time, is a record of the design and implementation of the AVIE, including its software, alongside a brief analysis of its various strengths and weaknesses.

The third significant contribution of this thesis is the artwork *La Dispersion Du Fils*. The work is unique in a number of ways. The work is the result of a collaboration with the Jean Michel Bruyère and the LFKs. Having spent a decade constructing an intricate and densely populated virtual world around the myth of Actaeon, Bruyère and LFKs bring to the work a byzantine world of images, characters and stories, music, film, text and sculpture, all of which provide the artistic foundation for the work and serve as a reservoir of source material. For this reason, the work possesses a richness and complexity of imagery that can only be attained through many years of creative practice. This, combined with the permissive framework of the doctorate allowing the work to be developed over a period of many years, results in a work with a degree of 'depth' that could be considered unusual in media art.

Technically speaking, *La Dispersion Du Fils* comprises an unusual, if not singular, synthesis of technologies and algorithms, which, when presented in the AVIE, results in an experience that can be described as unique.

1.4 THESIS STRUCTURE

This text is divided into three major components. The first presents a theoretical account of presence, the second concerns the AVIE, while the third is dedicated to the artwork *La Dispersion Du Fils.*

As the work is highly inter-disciplinary, reviews of literature and previous work are not presented in a single preliminary chapter, but distributed throughout the text, where they may best address the topic at hand.

2. PRESENCE, IMMERSION AND THE PANORAMA

2.1 THE PANORAMA AS VIRTUAL REALITY

In 1787, in the patent awarded for his invention "La Nature à Coup d'Oeil," Robert Barker wrote that the intended purpose of his apparatus was "to make observers, on whatever situation he may wish they should imagine themselves, feel as if really on the very Spot" (1796). The apparatus in question was the panorama, or cyclorama as it would later be known. Enormous cylindrical paintings, carefully lit, decorated and housed in elaborate purpose-built pavilions, Barker's panoramas apparently enjoyed great success, proliferating across Europe, North America and as far afield as Australia.[7]

A little over 100 years later, as panoramic paintings were "ceding to the taste of the day" and being, one by one, "converted into a circus, another into a skating rink and still another into a bicycle track" (Scientific American, 1896, p. 120), Charles A. Chase's *Electric Cyclorama* would breathe new life into the format by drawing on "the most recent progress and discoveries in the way of panoramic photography, projection apparatus, electric lighting, kinetoscopes, kinematographs and all other systems that permit of faithfully representing the phenomena of motion and life, as well as landscapes and views of inanimate objects" (Scientific American, 1896, p. 120). In the first of three patents awarded for his contraption, Chase writes how

> everything in view from the point where the photograph is taken will be reproduced exactly as it appears when seen from such point. It will thus be seen that landscapes from all parts of the world can be reproduced so that the spectators may see them as they would appear when seen in reality, and that the interior of the buildings maybe reproduced so as to appear to the spectators as seen from within. By this manner of reproducing views a person can get a better idea of the different parts of the world without actually going there than in any other manner heretofore devised. In fact he may see such views exactly as they would appear if seen on the ground. (Chase, 1895a, p. 2)

[7] Oliver Grau (1999) estimates that between 1870 and 1900 alone, 300 to 400 panoramas in Europe and America were enjoyed by over 100 million people.

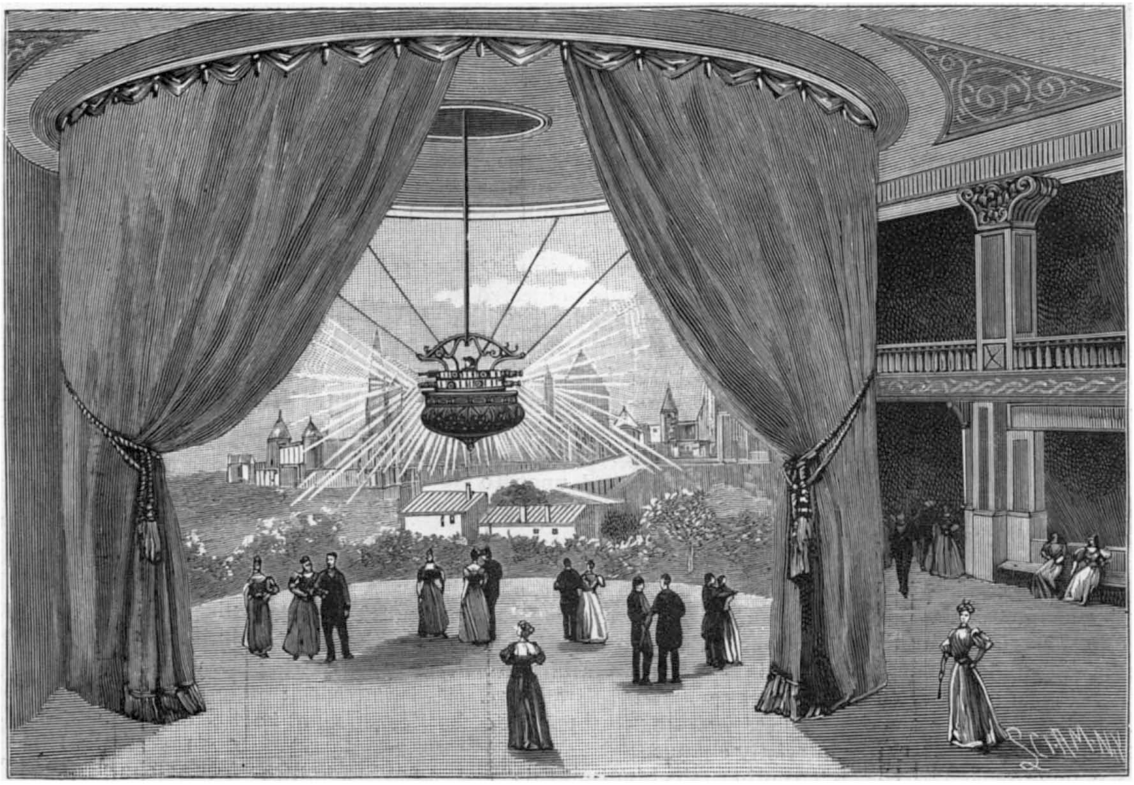


Figure 1 - Charles A. Chase's Electric Cyclorama 1895, illustrated in Scientific American (1896, p.120).

Five years later, Raoul Grimoin-Sanson would exhibit his Cinéorama at the Paris World Exposition of 1900. Having filmed a flight in a hot-air balloon with a ring of 10 synchronised cinema cameras, Grimoin-Sanson, in what was perhaps the world's first panoramic cinema, projected the resulting films as a single seamless panoramic image around a replica of the balloon, complete with a gondola, within which the audience stood. According to the patent granted Grimoin-Sanson in 1896, his *Multiplex Projector* was intended to "record and project moving panoramic views, giving the impression of reality" (Grimoin-Sanson, 1896b).

Grimoin-Sanson was not alone in his pursuit of a panoramic cinema. His compatriot and contemporary, Auguste Blaise Baron, filed a patent in 1899 for an even more ambitious mimetic device; a "device for circular, panoramic, animated projections in colour and with sound, known as the talking Cinématorama" (Baron, 1899). Baron explains:

> The aim of this invention is to have spectators travel all over the world without tiring [...]. In the operation of the device, the spectators will occupy the central

> part of the circumference described by the screen; they will be able to sit down or walk around and, as a result, be able to imagine themselves right inside the projected city. Baron (1899) translated in Mercier (1998).

It seems that the intended purpose of all these panoramic devices was to provoke in the audience a sense of "being there," or, as we might say today, a sense of *presence.* In this respect, they are all virtual reality apparatus, conceived with the technologies of the 18th and 19th centuries. Arguably, this ambition to induce presence can be traced throughout the history of panoramic displays, from the aforementioned experiments in panoramic cinema at the turn of the 20th century, to the amusement parks and world fairs of the 1960s, on to the experimental *Expanded Cinema* of the 1970s such as Stan VanDerBeek's *Movie-Drome*, and on to the most recent digital incarnations, such as Jeffrey Shaw's dome and cylinder displays, or Michael Naimark's *Be Now Here,* of which he too states was created in pursuit of a mediated experience "just like being there" (Naimark, 2005).

That the panoramic form should be repeatedly employed in the pursuit of presence seems, at first, obvious; the panorama permits an image to "escape its frame" and, by encircling the viewer, places them "within" the depicted scene.

But what exactly is this relationship between the panorama and presence? What is meant by this vague notion of being "within" an image? What perceptual mechanisms are at work? And why, even when the viewer is completely encircled by an image, do they seldom, if ever, feel truly present in the landscape of a panorama? Most importantly, what exactly does it mean to be immersed or feel presence?

2.2 INTRODUCTION TO IMMERSION AND PRESENCE

This thesis begins with immersion and presence, the twin phenomena at the heart of the notion of "being there."[8] The point of departure is the general sense of immersion that is seemingly a common feature of all media, from literature to cinema; from theatre to computer games. It is suggested that any common underlying mechanisms shared by these different manifestations of immersion, if indeed there are any at all, are outweighed by their differences, and that the term immersion is a slightly misleading polyseme for a wide variety of different phenomena. As such, subsequent discussion is constrained to the types of immersion arising from visual and aural representations of virtual worlds, the study of which has come to be known as "presence research."

Following this, an introduction to some of the key issues and stances in "presence research" is provided, and the distinction between immersion and presence is made clear. This provides the background for the next section, in which a novel framework for reasoning about immersion and presence is presented. The purpose of this framework is to make explicit the assumptions underlying, and the ramifications arising from, the *presence hypothesis* – the very idea that a sense of 'being there' in an *imperfectly* mediated world might be possible.

Once this framework is in place, discussion turns to perception of light. Here, the relationship between presence and perception of the environment is made explicit, when presence is defined as *active, coherent, multi-modal perception of invariants*.

Armed with this precise definition of presence, the role of the image in immersive media is considered, and it is argued that it is the use of images to convey a field of light that is most damaging to presence. It is then shown that the panorama represents a compromise between a mediated experience that is shared, and one that is immersive.

The proposed model of presence also sheds light on the role of interaction in the stimulation of presence, where it is recognised as important only insofar as it aids active-perception of the environment.

---

[8] The terms *immersion* and *presence* will be precisely defined in Section 2.2.2 below. Until then, the terms will be used interchangeably to denote a general sense of "being there."

Finally, it is demonstrated that the understanding of presence arrived at here presents a useful tool for the analysis of many different forms of mediated experiences.

2.2.1 A General Sense of Immersion

Immersion, in the general sense of "being there" or within some form of mediated experience, is a familiar concept. Familiar in the sense that it seems to have been a persistent facet of the arts since antiquity, and familiar in that it is experienced by many of us in one form or another on a regular basis. When it occurs it is easily recognised, and that we are somehow immersed when we watch a film, visit the theatre, listen to a radio play, read a book, or play a computer game seems somehow obvious and true. And yet, at the same time, immersion remains something abstruse and fleeting. It is strictly subjective, private, accessible by introspection alone and even then only barely, as if extinguished by our inward gaze. We may know it intuitively, but attempts to formalise our intuitions into rigorous theory remain highly speculative, fractious and contentious. The path from a general notion of immersion to a more precise and rigorous theory of immersion is not, it seems, without significant obstacles.

What might a theory of immersion look like? It would provide a clear definition, allowing demarcation of what is and is not immersion, and between any variants of immersion, should they exist. It would furnish us with a vocabulary and taxonomy free from ambiguity. It would explain the passage of information from virtual representation to physical stimuli to perception to conscious experience, couched coherently within some theory of mind or perception. It would be predictive, and prescriptive in a manner that is beneficial to the creation of immersive experiences. It would serve as a means of understanding existing immersive creations, and as a foundation for the creation of new ones. It might also prescribe a means of measuring immersion, or in the very least explain why such measurement is not possible.

Many of these theoretical elements can be found in one form or another in scholarly discourse on all forms of media, from literature, theatre, cinema and television and the visual arts to computer games and simulation, where immersion has been described or studied under such names as "suspension of disbelief" or "dramatic illusion"

(Coleridge, 1817) "ideal presence" (Kames, 1765), "transportation" (Gerrig, 1993), "absorption" (Cohen, 2001), "immediacy of mediated experience" (Bolter & Grusin, 1999), "engagement" (Laurel, 1991), "aesthetic illusion" or "experiential illusion" (Wolf, 2008), "involvement" and "psychological participation" (Walton, 1990), "illusion of reality" (Gombrich, 1960), "effet de réel" or "illusion référentielle" (Barthes, 1968) and "flow" (Csikszentmihalyi, 1990), "telepresence" (Minsky, 1980) or simply "presence" (Negroponte et al., 1980).[9]

Given the diversity of media within which mention of immersion arises, it is natural to ask whether all or any of the properties of immersion, and the mechanisms that give rise to it, are common across media. After all, in everyday language the single term 'immersion' can be used in reference to a book, film or video game meaningfully and without confusion. Might we not seek a theory of immersion that somehow simultaneously encompasses all these different contexts? The allure of a "pan-medial" theory of immersion is difficult to resist partly because there exists a continuum between all the different media that might give rise to immersion - literature, text, spoken word, video games, simulation, cinema -, such that it is a very simple exercise to conjure examples of mediated experiences that fall arbitrarily between any of these archetypical categories. Therefore, any theory of immersion tailored for a specific medium or context - no matter how narrowly it focuses - will always be confronted with fringe cases, which, should they be accommodated, only present more fringe cases. Like this, there is an irresistible pull towards a single, nomological theory capable of accounting for immersion in all its manifestations. Werner Wolf, in writing of aesthetic illusion (a common term for immersion in literature), concludes:

> Owing to the dependency of immersion on the semiotic macro-frames of narrative and description as well as on the media and the genres used, a desideratum for future research is certainly interdisciplinary cooperation, not only between narratologists and cognitive psychologists, but also, and closer to aesthetic concerns, between narratology and drama theory, art history and film studies. For aesthetic illusion is a transmedial, transmodal and transgeneric

[9] For this list of terms I am partially indebted to Werner Wolf (2013).

> phenomenon, and if this is taken into account, a still better understanding of it will be achieved, ultimately leading, perhaps, to a general theory of aesthetic illusion that transcends individual genres, modes of representation and media. (Wolf, 2013).

Wolf's writings on aesthetic illusion offer a fine example of such a "transmedial, transmodal and transgeneric" approach to immersion. He defines aesthetic illusion as "a feeling, with variable intensity, of being imaginatively and emotionally immersed in a represented world and of experiencing this world in a way similar (but not identical) to real life" (2013). Defined as such, Wolf suggests that aesthetic illusion can be elicited by "narrative fiction, drama, lyric poetry, painting, sculpture, photography, film, and contemporary virtual realities such as computer games, while excluding (most) instrumental music" (2013). In compiling a list of some "typical characteristics of illusionist representations and principles of illusion-making" found in fictional textual narratives, Wolf (2008) identifies some features of immersion that could be considered universal:

- The principle of consistency; the illusion that the virtual world exists independently of both the participant and the mediation, and that this world obeys some kind of inherent and intrinsic logic.
- The principle of life-like perspective; experience of the virtual representation in a manner commensurate with everyday experience.
- The principle of *celare artem*;[10] to conceal the means by which art is achieved, leading ultimately to the *disappearance of the medium* or the *illusion of non-mediation*.

These principles appear to be applicable to any medium, which allows us to conclude that at least some aspects of immersion are universal.

Another approach to a "transmedial" treatment of immersion is found in the application of concepts originating from virtual reality to traditional non-digital media.

[10] *ars est celare artem*: "the true art is to conceal the art," a maxim often attributed (seemingly erroneously) to Ovid's *Ars Amatoria*.

Marie-Laure Ryan (2001), for example, explores the implications of viewing literary narratives as forms of virtual reality. In fact, the *ideal* of virtual reality has had such an influence on concepts of immersion and mediation in all domains, that writings and thoughts on immersion in any domain can be divided into those that precede the invention of virtual reality, and those that follow. Post-VR discourse on immersion, even in non-digital media, has been greatly influenced by such ideas as virtuality and virtual worlds, simulation and algorithms, interfaces and interaction, agents and agency, and the very possibility of virtual worlds existing independently of representation, mediation or real-world counterparts. Virtual reality, despite itself not yet actually existing (a virtual virtual reality, so to speak), has rendered these concepts somehow more tangible and vital, and in doing so irrevocably changed the way we see all media, both past and present. Virtual reality has given the question of immersion in any medium a certain sense of weight and urgency, by transforming it from a curious and somewhat nebulous side-effect of fiction, to a very concrete, real and achievable end in itself. It is this post-VR version of immersion that Oliver Grau is thinking of when he writes that immersion "is undoubtedly key to any understanding of the development of the media" (2003, p. 13), before going on to show that the history of panoramic images, from the frescoes of ancient Rome to the IMAX of today, might all be explained within the framework of virtual reality.

The promise of a transmedial theory of immersion is that an understanding of some aspect of immersion gained from one context or medium can be applied to others. In some respects this is certainly the case. For example, an understanding of how one might be immersed in a narrative will likely apply to all narrative media, from radio plays to literature.

However, care must be taken to avoid drawing too much from superficial analogies. A sense of "being there" might arise in all media, but the sensations and mechanisms in each case may be utterly different. The "being there" provided by a book is, upon reflection, not the same as that elicited by a VR system such as the CAVE.[11] The

[11] The CAVE system will serve often throughout this text as a canonical example of a virtual reality system. The reader is referred to Cruz-Neira et al. (1992) for more information.

difference may not be just a matter of degree, for the two instances of "being there" may have nothing in common beyond the words we use to describe them. The language of immersion, it seems, is highly polysemic, and the same terms used in relation to different media, describe very different phenomena. Further, there are mechanisms at work in the CAVE that simply have no analogue in text. Take for example vection, the illusion of self-motion arising from visual stimuli that has been demonstrated to play a part in spatial presence (Riecke et al., 2011). Seeking an analogue of vection in literature would certainly be an illuminating exercise, but it is difficult to see how anything but a metaphoric interpretation of vection may play a role in any theory of literary immersion.[12]

In fact, as we ask whether former concepts such as "aesthetic illusion" or "transportation" or "suspension of disbelief" might all be captured by a single concept, immersion itself is splintering into finer classifications. It is increasingly clear that broad definitions of immersion such as Wolf's "being imaginatively and emotionally immersed in a represented world and of experiencing this world in a way similar […] to real life," are in fact short-hand "catch-alls" for a complex array of distinct and varied phenomena such as spatial presence, sense of realism, narrative immersion, sense of social presence, sense of other and sense of agency.[13] Each of these varieties of immersion may have different causes and properties in different media or contexts, resulting in a picture of 'immersion' that is highly fractured and the very antithesis of a unified "transmedial, transmodal and transgeneric" theory.

For these reasons, transmediality will not be held as a necessary property of any definition of immersion or presence. Rather, from this point on, discussion of immersion and presence will be restricted to a very particular domain: visually and

---

[12] I do not discount the possibility of a theory of literary immersion that, by drawing on notions of embodied cognition such as those proposed by Mark Johnson finds some analogue of vection in linguistic thought (Johnson, 1987). Nor the possibility of neural correlates between reading and understanding the phrase "moving forward" and physically moving forward. But it is certainly not obvious how, even with such connections, knowledge obtained from psychophysical experiments in vection can be applied to text.

[13] For a survey and literature review for all these different varieties of immersion see Lombard and Ditton (1997).

aurally mediated virtual environments. The study of immersion and presence in this context has come to be known as "presence research", the subject of the next section.

2.2.2 An Introduction to Presence Research

Over the last two decades, study of immersion in virtual environments, such as those found in computer games and virtual reality, has blossomed into a research domain in its own right.[14] This body of research, which has come to be collectively known as "presence research," is interdisciplinary in nature, drawing together psychology and psychophysics, cognitive science, computer science and engineering. The results of the research, which are ostensibly an understanding of the causes of presence, have found application in the design of computer games, simulations for training or treatment of phobia or pain relief, education, social robotics, telepresence, user interface design and philosophy, to name just a few.

Presented here is a brief introduction to various aspects of presence research. Rather than attempt a comprehensive survey of the field (for far more comprehensive accounts than could possibly be provided here can been be found in Lombard and Ditton (1997), Lee (2004) and Lombard and Jones (2006)), the focus here will be on a selection of the most important elements that distinguish the various definitions or models of presence, as found in presence literature. This proves an efficacious means of gaining an understanding of the breadth and depth of presence research, and lays the foundation for the development of our own picture of presence, which takes place in Section 2.4 below.

Presence from Immersion

The terms 'immersion' and 'presence' are used here in the manner proposed by Mel Slater (2003). To be immersed is to perceive and interact with some virtual environment, entity or narrative through some medium. Presence is a certain experiential quality or

[14] Presence research is not exclusively concerned with immersion in virtual environments, but also includes interactions with robots and intelligent virtual agents, as well as video conferencing and telepresence. Again see Lombard and Ditton (1997) for a discussion of presence research that includes all these notions.

aspect of being that arises as a consequence of this mediation. Immersion is a property of the medium, and presence is a state-of-being in the viewer. The more a medium is able to instil a sense of presence in the viewer, the more immersive it can be said to be. Exactly what constitutes an immersive system, or what determines the level of immersiveness of a system is, therefore, entirely dependent on how one defines presence.

To repeat: to be immersed is to experience a virtual environment through some form of mediation. To be present, or feel presence, is to possess some particular aspect of being in response to this mediation. The exact nature of this aspect of being, which may be a feeling or psychological state, or a type of relationship with the virtual environment, depends on one's definition of presence. It is here that the notion of presence splinters into a multitude of different subspecies or "dimensions": spatial presence, sense of realism, physical presence and sense of embodiment, sense of social presence, sense of other (co-presence) and sense of agency are some examples. They all reflect some quality of being-in-the-world, and correspond to such utterances as "I felt like I was there," (spatial presence), "but it didn't feel necessarily real - more like being trapped in a cartoon" ((lack of) sense of realism), or "it seemed natural" (plausibility), or "I felt present, but invisible, like a ghost" (spatial presence, with lack of embodiment), or "I felt constrained in my capacity to act" (lack of sense of agency), "I felt like I was in the company of other minds" (social presence or co-presence), and so on. These various notions often overlap, and a general sense of presence in a virtual environment possessing a degree of richness and complexity found in the real world (with people, stories, places, events) is sure to involve a combination of all these different dimensions of presence. Which of these different dimensions of presence are more or less important to an overall sense of presence, if such a thing can be imagined, depends entirely on context. Social presence, for example, is irrelevant to a virtual environment without characters, but pivotal to an immersive "chat-room" experience.

Presence in the Real World?

The first question faced when deciding upon a definition of presence is whether or not presence is something we experience in the real world. On one hand, presence may be taken to be something that arises naturally in everyday experience of the real world, rather than some alien state or quality of being particular to mediated experiences. In this case, the presence that occurs in the real world is implicitly the standard by which we measure presence in a mediated environment. There are implications to this. If we are seeking to replicate some natural phenomenon that occurs in the real world, we must surely observe it and understand it there, in its natural habitat, before trying to replicate it artificially. For example, is it possible to *not* be present in the real-world, or is it a constant feature of existence? Must one always be present somewhere? Can one exhibit degrees of presence, or is it a binary on-off arrangement? Is it possible to attain states of "hyper-presence," above and beyond that which might be attained in the real world? Intriguingly, it may be that only by studying presence in mediated environments that an appreciation of these aspects of presence in the real world is attained. Example definitions of this form of presence include "the intuitive perception of successfully transforming intentions into action" (Riva et al., 2007) or "a feedback from unconscious cognitive processes that informs conscious thought about the state of the spatial cognitive system" (Schubert, 2009), both of which are definitions that make no reference to reality or virtuality, and so apply equally to being in a real or virtual world.

On the other hand, presence may be defined as something special and unique to mediated experiences. One example of such a definition is that proposed by Lee (2004), where presence is "a psychological state in which virtual objects are experienced as actual objects in either sensory or non-sensory ways." Another is that provided by International Society for Presence Research:

> Presence is a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience. (ISPR, 2000)

Both these definitions constrain presence to mediated experiences. This is, however, only a superficial constraint, for both definitions *imply* replication of some aspect of being in the real world. They both imply that the psychological state identified as presence is first and foremost a real-world phenomenon, but choose to reserve the term 'presence' solely for interactions with media. Note also that both definitions amount to little more than a restating of the "disappearance of the medium," (Lombard & Ditton, 1997) and say very little about the nature of presence itself.

A more intriguing prospect is the identification of presence with states-of-being that are truly peculiar to mediated experiences, and have no counterpart in the real world. For example, presence might be defined as arising solely from immersion in an *imperfect* simulation. Or perhaps *knowing* that the world is an illusion is a necessary ingredient for presence. It might be that we never actually feel presence, but *absence*; the feeling of being somehow only partially immersed within a world.

Sense of Presence

Presence theories could also be divided into those that see presence as entirely seated in consciousness, or as subconscious phenomena sometimes accompanied by conscious phenomena, and those that are not directly visible to the conscious mind at all, but must be detected indirectly. For example, Schubert suggests that spatial presence is a *feeling* fed back from

> […] unconscious processes of spatial perception that try to locate the human body in relation to its environment, and to determine possible interactions with it. If the spatial cognition processes are successfully able to locate the body in relation to the perceived environment, and construct possible actions in it, the feeling of spatial presence is fed back and becomes available for conscious processes. (Schubert, 2009)

In contrast, Lee (2004) defines physical presence as "a psychological state in which virtual (para-authentic or artificial) physical objects are experienced as actual physical objects in either sensory or nonsensory ways." Here, conscious sensation of presence is not required (although the term "experienced" does add a certain ambiguity here). In general most presence theories define presence as either being, or being accompanied by, some kind of conscious sensation; a feeling or "sense of…" something. However, as will be seen, this need not necessarily be the case.

Emotions and Presence

Another distinguishing feature among definitions of presence is the role attributed to emotion. On one hand, emotions are held to be "orthogonal" to presence, which is to say that a viewer's level of presence is not married to their emotional state, and whether we feel afraid, nervous, happy or excited does not in of itself make us more or less present. Consider, for example, presence as "the extent to which people respond realistically within a virtual environment, where response is taken at every level from low-level physiological to high-level emotional and behavioural responses" (Sanchez-Vives & Slater, 2005). The consequence of such a definition is that presence *cannot* be linked to any particular emotion, for whether we are immersed in the midst of a great battle, or in a laundromat waiting for our laundry to dry, we may be equally present.

This does not, however, preclude emotional response to an immersive experience being taken as a symptom or measure of presence. For example, failure to respond emotionally to a highly evocative virtual scenario may be evidence of lack of presence. Indeed, experiments that place participants in various virtual precarious predicaments, such as standing on the edge of a virtual cliff, demonstrate that anxiety and fear are measurably heightened when presence is strongest (Sanchez-Vives & Slater, 2005).

Further, it is possible that we are more or less sensitive to the *artefacts* of mediation in different emotional states, for if emotions can alter perception, then they can alter the way we see the flaws in a mediation, thereby directly influencing the level of presence. The "willing suspension of disbelief" or the complicity of the viewer could be considered examples of this.

On the other hand, there are the definitions of presence that willingly admit the possibility of emotions directly influencing presence. And it seems that, once such a definition is adopted, experimental evidence to support such a case is at hand. For example, Riva et al. (2007) suggest that sense of presence is greater in "emotional" environments. Further, the connection between emotion and presence was found to be bi-directional. Experiments "showed a circular interaction between presence and emotions: on one side, the feeling of presence was greater in the "emotional" environments; on the other side, the emotional state was influenced by the level of presence" (Riva et al., 2007, p. 45).

Aspects of Presence Research

Presence research is generally marked by a desire to isolate the 'essence' of presence, which is to reduce the requisite aspects-of-being to some kind of minimum set, beyond which sense of presence is lost. There is a further desire to isolate the aspects of presence that hold true for all environments and people. Whether or not there are such 'universals' of presence is one of the questions facing presence research.

Parallel to identifying different types of presence and proposing cognitive or psychological explanations, presence research is also concerned with the exact relationship between immersion and presence. This is the study of how all the different variables of mediation, such as image size, frame rate, resolution and stereoscopy (Ijsselsteijn et al., 2001), or visual realism (Slater, 2003; Slater et al., 2009) influence presence. See Cummings et al. (2012) for a recent attempt to catalogue such studies, dating back to 1995.

Presence research also involves identifying certain physical stimuli or psychophysical cues that have a particular strong correlation with presence. For example, *vection* - the illusion of self-motion - has been shown to increase spatial presence (Riecke & Schulte-Pelkum, 2013). In turn, spatial sound has been shown to enhance vection. Binding of multi-sensory cues (Harvey & Sanchez-Vives, 2005) and causal perception – the direct perception of cause and effect (Cavazza et al., 2007) - have also been shown to influence presence. Such knowledge can be employed in the design of virtual environments, by deliberately creating worlds rich with these kinds of sensory cues.

Presence research is also concerned with the problem of measuring presence. A great deal of approaches can be found in literature, ranging from plain introspection –

i.e. simply asking the participant to report on their perceived level of presence (Witmer & Singer, 1998) – to physiological cues such as skin conductance or cardiac rhythm (Wiederhold et al., 2001); or observation of involuntary behaviours such as startle reflex or posture (Freeman et al., 2000; Lepecq et al., 2009) to measurement of neurological activity using such technologies as EEG (Kober et al., 2012) or transcranial Doppler sonography (Rey et al., 2010).

Conclusion

The intention of this brief introduction to the field of presence research was to impart some notion of the depth and breadth of research into the phenomena of immersion and presence. It has been seen that, even when the focus is restricted to immersion within interactive virtual environments, a number of different forms of presence can be identified, and for each of these a variety of explicatory models have been proposed. Any discussion concerning presence or immersion must therefore begin with precise definitions of the terms, so that their intended meanings can be distinguished from the numerous possibilities on offer. It is to this task that this thesis turns in the following sections.

2.3 MIMETIC IMMERSION

To be immersed is to perceive and interact with a virtual environment, entity or narrative, through some form of mediation. To be present, or feel presence, is to possess some aspect of being in the world that one would normally have, were one to experience this environment, entity or narrative directly, without mediation. This approach to presence could be called *mimetic immersion*, or more simply *mimesis*, not because it requires imitation or simulation of the real-world (although, as will be seen, this is one avenue towards presence), but because it is predicated on the replication of some aspect of our being or relationship with the real-world.

Before tackling the question of what this aspect of being might be, it is important to discuss certain implications arising from this mimetic approach to presence. In order to do so, I will introduce a short-hand to aid explanation. Mimetic immersion has the following elements:

I - *Immersant:* The conscious participant in whom we hope to excite the sensation of immersion. This is the *user, player, reader, viewer, participant* or *audience member*, but as all these terms allude to higher or lower amounts of passivity or interactivity, the deliberately neutral term *immersant* is preferred.

E - *Environment*: The 'real' world. The entity, environment, event, narrative, place or phenomenon within which the immersant is to be immersed. This may or may not be fictional.

V_E - *Virtual Environment*: A virtual representation of E.

M - *Mediation:* Some method of mediating to the immersant the digital representation V_E using one or more physical carriers of information (light, sound, force) to stimulate one or more human senses (vision, hearing, touch), which we write $M(V_E)$. Note that the line between M and V_E is arbitrary, and sometimes the separation of M from V_E is meaningless. $M(V_E)$ and E are of equivalent class, meaning they can be compared, and it is meaningful to talk about the fidelity with which $M(V_E)$ emulates E.

S and S' - *State-of-being:* The mental and physical state of the immersant. This notion of state-of-being can be taken as a complete description of all the aspects of being-in-the-world that might vary during the course of a mediated experience, including such things as moods, thoughts, knowledge, memories as well as physical posture and motion, breathing and heart-rate. No distinction is drawn between conscious and unconscious aspects of being, as there is nothing to suggest that presence is limited to conscious processes alone. The mediated state S' is the state-of-being of the immersant arising from exposure to the mediated environment $M(V_E)$, while the immediate state S arises from direct interaction with the real world E. S is a function of the immersant I and the environment E which we write as $S(I, E)$, where S' is a function of the immersant I and the mediation $M(V_E)$, and written as $S'(I, M(V_E))$. It is noted that mental phenomena are temporal in nature, and it makes little sense to talk about states of being without reference to time, but for the discussion that follows here, an atemporal formula is sufficient.

Using this crude algebra,[15] the immersant I is guaranteed to be present when their mediated state S' is identical to the unmediated state S, which we might write as:

$$S'(I, M(V_E)) = S(I, E) \qquad \textit{Eq. 1 Mimetic Immersion}$$

If S' = S, and S encapsulates presence, S' must also. This has two immediate implications. First, there is no "state of presence" *per se.* Rather, there is simply the state the immersant would have, were they to experience E directly. It is simply nonsensical to talk about the level of presence in S' without recourse to S, I and E.

Second, this equation immediately suggests an avenue for ensuring presence in a virtual world, for it implies that presence is guaranteed if the mediation of the virtual world is physically identical to the real world. That is, if $M(V_E) = E$, then $S' = S$. This can be considered the naive or 'brute-force' approach to mimesis, which we might call *physical mimesis*, for it demands a complete physical replica of E. In it is the implicit assumption that the replication of the physical stimuli (forces, light, sound) arising from an environment is sufficient to generate presence, and that there are not some properties of the world that defy physical mimesis, and yet are required for presence.

Physical mimesis is a hypothetical ideal, a declaration of possibility of arriving at presence through mimesis. A more realisable approach is to note that $M(V_E)$ need only be *indistinguishable* from E *to the immersant*. By taking into account the acuity and domain of our sensory interactions with the world, the complexity of the physical stimuli needed to emulate E can be reduced. For example, rather than replicating the complete sonic wave-field, the wave-front need only be reproduced at the recipient's ears.[16] A more subtle approach is to exploit features of perception where very different stimuli are perceived as identical. Consider the manner in which the human eye reduces an arbitrary spectrum of wavelengths to just three chromatic intensities (corresponding

---

[15] I have introduced the symbols I, E, $M(V_E)$ and S for the sake of brevity only. The equations that follow are intended only as short-hand for their natural language equivalents, and should be read that way. I am not in any way proposing a mathematical relationship.

[16] A technique sometimes referred to as binaural synthesis. See, for example Kuhlen et al. (2007).

roughly to red, green and blue). Thanks to this colour *metamerism*, as it is known, light composed of just three suitably chosen wavelengths is perceptually indistinguishable from light composed of an arbitrary spectrum of wavelengths. A mediation, therefore, need only reproduce light stimuli at these three specific wavelengths rather than emulate the full spectrum of the real world counterpart.[17] Another example is *summing localisation*, where two coherent sounds emanating from two different locations can, under certain specific conditions, be heard as a single 'phantom' sound emanating from a single position in space between the two (Warncke, 1941; Blauert, 1997). As metameric stimuli are *perceived as identical*, the resulting states-of-being are also identical, and so preservation of presence is ensured. This approach could be called *metameric mimesis.*

The pursuit of metameric mimesis calls on science of perception and psychophysics, science and engineering of display technology (visual, aural, haptic etc.), and the computer science domains of graphic, synthesis and simulation. It should be noted that metameric mimesis is still predicated on S' = S, so the mediated world must *appear identical* to E. A pixelated view of V_E is *not* indistinguishable from E if the pixels are visible, so this does not qualify. Given the current state of immersive technology, metameric immersion remains an ideal not yet attained.

2.3.1 The Presence Hypothesis

Until this point, there has been no mention of which aspects of our state-of-being S must be faithfully reproduced in order that presence is guaranteed to be preserved. It has simply been reasoned that if the mediated state S' is identical in every way to S, and S exhibits presence, then S' must do it as well. And it may be that this is a necessary requirement for presence, that S' = S. It may be, however, that there are some aspects of being that are irrelevant to presence. This possibility – that there are some aspects of being that are salient to presence and others not – we might call the *presence hypothesis.*

One piece of evidence supporting the presence hypothesis is that unless the viewer has no knowledge whatsoever that they have entered a virtual environment, then S' can

[17] This convenient feature of vision forms the basis for nearly all colour image sensing, storage and display technologies, from photography to television, printing, and digital imaging.

never be identical to S, because only S' incorporates knowledge of being in a mediated environment, while S does not. But anyone who has ever reported feeling a sense of presence in a virtual environment has undoubtedly been aware that they are participating in a simulation, so knowledge or ignorance of the mediation cannot be a relevant feature of presence. Further, the knowledge and novelty of being in a virtual environment often induces an emotional state (excitement or anxiety, for example) that would not be experienced in the equivalent real world situation. Again, observations suggest that such deviations in state do not diminish presence.[18]

Deciding which properties of being in the real world must be faithfully reproduced in the virtual world in order that presence be preserved is equivalent to defining some measure of distance between S' and S that gives weight to the salient features of S and ignores others.[19] This *measure-of-presence* might be written as $\Delta P(S', S)$ which allows the presence equation to be rewritten as:

$$\Delta P\left(S'(I, M(V_E)), S(I, E)\right) = 0 \qquad \textit{Eq. 2 Measure of Presence}$$

The presence hypothesis allows for the preservation of presence when S' is not identical to S, and in doing so allows for presence when $M(V_E)$ is not perceptibly indistinguishable from E. It allows for the kinds of incomplete or imperfect mediations that are attained with current immersive technologies. For example, that S' differs from S because the virtual world is clearly pixelated, is no longer a necessary impediment to presence. This extends to all matters of fidelity and display artefacts, such as frame rate, colour gamut, resolution, field of view and so on. Further, it implies that the mediation

---

[18] This should not be understood as a proof – for it would be circular – but an observation. Sightings of presence in the wild are the main reason for believing the presence hypothesis to be true.

[19] Another way to conceptualise the presence hypothesis is to imagine states S and S' as points in the space of all possible states-of-being. The presence hypothesis is that presence persists not only when $S' = S$, but when S' falls in some region of space around S. The shape of this region depends on the definition of presence and which dimensions of being are considered more or less salient to presence. We might write the presence hypothesis as:

$$S'(I, M(V_E)) \in \boldsymbol{S}(I, E), \qquad \text{where } \mathbf{S} \text{ is a set of states in which presence is preserved.}$$

Deciding which properties of being in the real world must be faithfully reproduced in the virtual world in order that presence is preserved is equivalent to defining the set **S.**

can be reduced to the 'bare essentials' required for presence. In the same manner that a moving pattern of dots can provoke strong sensations of being in motion when their movements result in the same optical flow that would result from actually moving (a well-documented psychophysical phenomena),[20] it may be possible to strip the virtual world down to some sparse collection of stimuli, and still maintain presence.

But more than this, it permits mediations that are markedly and purposively different from E. It allows for $M(V_E)$ that differ from E not because of limitations in technology, or failure to emulate the real world, but are intentionally different, such as abstract or stylistic interpretations. Features of the world may be exaggerated or understated. It allows the addition of non-diegetic elements; perceptual cues that would not be present in the real world, such as music or narration. In short, the presence hypothesis allows for the development of an artistic practice of presence and immersion, and opens up the possibility of immersive *poesis*.

2.3.2 The Universal Presence Hypothesis

Now, for a particular environment, it might be that whether or not the immersant feels warm or cold has no bearing on their presence in E, and so sense of warmth or cold can be omitted from the measure-of-presence ΔP. It is, however, an easy exercise to conceive of a particular situation in which feeling warm or cold might be considered important to presence. For example, an immersant may never report feeling present in a simulation of the arctic if they feel warm. Similarly, for a class of environments where the viewer never moves, sensations of ego-motion can be omitted from the equation. But once the viewer moves, ego-motion becomes an important determinant in presence. Does this imply that aspects-of-being salient to presence change with each environment? There is no reason why this should not be the case, and this framework allows for this possibility.

A similar source of variation is the immersant themselves. The explicit inclusion of I in the equation entails the possibility that what may invoke presence for some, does not for others. Visual acuity, for example, varies for everyone. But more than this, it allows

[20] See for example Burr and Thompson (2011).

for definitions of presence that are sensitive to subjective, cultural and personal factors in the invocation of presence. Our relationship to technology, our previous personal and collective experiences, our culture, learned behaviours and skill, may all impact our capacity for presence. Further, our knowledge of and previous experience with the real-world E may affect our capacity for presence.

The presence hypothesis was previously defined as the existence of certain facets of being-in-the-world that are salient to presence for a particular environment and immersant while others are not. This is now extended to the hypothesis that there are certain facets of being-in-the-world that are salient to presence for *all possible immersants and environments*. This is to say that while there may be innumerable definitions of presence that depend on subjective (I) and contextual (E) parameters, the ones that are of interest are those that hold for all subjects and all contexts. Presence research is largely concerned with such universal concepts of presence that do not vary with subject or context. Note that all of the definitions of presence given in the introduction above are examples of universal hypotheses.

2.3.3 Simulation of Fictional Worlds

The most important implication of this stronger version of the presence hypothesis is that, by inductive reasoning, it permits mimetic immersion in *fictional worlds*. If "all possible E" is interpreted to entail fictional worlds, then the hypothesis implies that presence is neither restricted to the real world nor imitations of the real world, but may arise in any world, providing that certain conditions are met.[21]

We now have the possibility that E, the 'real-world' environment or entity being mediated, does not physically exist. It may be a fiction. It may only exist in other mediations (for example, a film based on a novel), or only in the imagination of the creator. Alternatively, it may only exist mathematically or algorithmically, such as a three dimensional fractal.

[21] It could be argued that the entire enterprise of computer games is predicated on, and verification of, this universal presence hypothesis.

If we consider mimesis as simulation, then we are presented with the intriguing prospect of simulating something that does not exist. At first, this appears absurd. Yet if we return to the canonical example of a simulation - the flight simulator - and imagine a faithful simulation of a flight from Sydney to New York, except that New York is now in New Zealand and Sydney is in Japan, is this not a simulation of a fictional world? The apparent absurdity arises from the different meanings of simulation, for the term, in the age of computers, has come to mean something more than simple imitation, and now implies a *process* or *model*. To simulate something is to build an abstract model of it and to then automate that model.

The concept of simulation not only as imitation but as process/model is intrinsic to mimetic immersion. Mimetic immersion is based on the transformation of the virtual world into physical stimuli perceptible to the senses, and these physical stimuli (sound, light or force, or whatever the channels of communication may be) must be specified for every permissible point-of-view in the virtual world. Further, a mimetic simulation must specify how this world evolves over time, and how the movements and actions of the immersant influence this evolution. These are the fundamental constraints placed on the types of abstract or fictional worlds amenable to mimetic immersion: they must prescribe the exact manner that world manifests as physical stimuli, and they must do so for all possible world-states and points-of-view. Theoretically this could be done by hand, but only by dramatically restricting the movements and actions of the participant, and thereby reducing the state-space of the simulation (including the position of the viewer) to a small number of possibilities. As these restrictions on the movements and actions of the viewer are removed, (and it will be shown in the following section that freedom from these restrictions is fundamental to presence), the amount of information demanded of the mediation becomes unbounded, making an entirely hand-crafted immersive mediation impossible. The only feasible solution is that these stimuli are the product of some automatic process or computation. And when these processes or computations employ some underlying representation or model of the world, they can be duly described as simulations.

Mimetic immersion is therefore predicated on a *processual* description of how the virtual world manifests *as physical stimuli*. This is tantamount to specifying a *physics* of the virtual world; a processual description of how a fictional or virtual world is ultimately represented and experienced as light, sound or force. The immediate question that follows is whether there are necessary or sufficient features that this virtual physics must possess in order that presence is sustained? Intuition would suggest that the closer the virtual physics matches the real world, the more likely that presence is preserved. This is the naive, brute-force approach introduced above as physical mimesis. But perhaps, just as there are aspects of being that are both salient and irrelevant to presence, there are both salient and irrelevant features of physics itself. Perhaps our sense of presence is resilient to certain aberrations or artifices. This is the case for visual realism, where the synthetic optics used to transform geometric models into computer-generated images is only a crude approximation of the real-world physics of light, and yet from these crude approximations photorealistic images can be obtained. In fact, the pursuit of photorealistic computer generated images can be seen not only as a progressively more accurate emulation of real-world optics, but as a cataloguing of features of optics that are more or less important to the illusion of realism than others. A similar search can now be undertaken with respect to presence, in which the features of the optics, haptics or acoustics that transform the virtual world into percepts are charted and prioritised with respect to their impact on presence. It may be that these virtual physics need share very little with the real world, beyond such properties as determinism, consistency and causality.

2.3.4 Observations

The definition of mimetic immersion given here asks us to imagine the state of being of the immersant were they to experience E first-hand, without mediation. Mimetic immersion, then, is asking us to imagine the fictitious state of being S arising from being in a fictitious world E. The ontological status of fictional worlds has long occupied philosophical debate and, while fascinating, it will not be entered into here save to say that if we can have fictional worlds then we can have fictional mind-states arising from these fictional worlds.

Putting aside the ontological puzzle of E and S, a more immediate problem is how would we know what S would be like? The short answer is that we cannot. Do we need to know? When we measure presence, either objectively or introspectively, we are necessarily comparing S' with S. We may, therefore, be asked to imagine such things as "what is it like to *really* be inside a three-dimensional fractal?" or "do I feel or behave as if I am *really* floating in an infinite field of cubes?" There seems no way of knowing, and so without access to S the comparison must instead be made with some more or less equivalent or *generic* experience in the real world.

A more subtle problem is that, given a representation $M(V_E)$ of a fictional world, and being asked to imagine its real world equivalent E and what it might be like to be in it, how do we distinguish between approximations and unintended flaws in $M(V_E)$, and features? For example, the virtual world might have been implemented using voxels, which manifest very visibly in the world as a pervasive cube-like structure. What is the real-world equivalent of this virtual world? Are the voxels intended to approximate a smooth surface, in which case the cube-like structure is an artefact of mediation and a potential distraction from presence? Or was the intention to simulate a cube-like world, in which case $M(V_E)$ is a perfect representation? It would seem that for definitions of presence that rely somehow on how plausible or 'real' the virtual world seems to the immersant, then indeed the same fictional world may give rise to different levels of presence according to whether the immersant sees elements of the mediation as errors or as features. This would be determined, at least in part, by presumptions on the part of the immersant as to the true nature of the virtual world, as well as their familiarity with the mediating technology, in which case flaws or artefacts of mediation may be more easily noticed. It would also depend on their familiarity with E, be it fictional or real, which leads to another question: is it easier or harder to immerse someone in a world that they know very well from reality, or from other mediations (a book for example)? Does our knowledge of the real world help us to overlook deviations, or omissions, in the mediation, or does knowledge of the real world make us more sensitive to such departures from reality?

2.3.5 Summary

In this section, mimetic immersion was introduced as a framework for reasoning about presence in mediated experiences. The framework posits presence as some aspect of being in the world that one would normally have, were one to experience the virtual environment, entity or narrative directly, without mediation. *This is not in itself a definition of presence*, but rather an analysis of the assumptions underlying the possibility of presence in a virtual environment. This work was undertaken following a review of presence literature, in which it became evident that many definitions or even theories of presence found in literature are developed without sufficiently addressing these underlying assumptions or their implications.

In the following section this framework is used to construct a working definition of spatial presence, with the goal of better understanding the sense of 'being there' that might manifest within a panorama.

2.4 Presence and Perception

Using the concept of mimetic immersion presented above, a precise definition of presence is now presented. It is derived from an understanding of how we perceive and interact with the real world. Presence is, therefore, taken to be a feature of our every day being in the world, and not at all unique to mediated experiences. Just as we never stop perceiving, presence is defined as a constant, unwavering feature of being in the real world, neither changing with event nor emotion. It is neither a feeling, nor a state, but a particular *relationship* with the world. It does not arise from this relationship, it simply *is* this relationship.

This understanding of presence is won by examining our natural manner of seeing and moving in the world. The point of departure, then, is the perception of light, and the twin concepts of the light field and the plenoptic function.

2.4.1 Light Fields and the Plenoptic Function

A *light field* is a conception of light as a dense field of points, through each of which innumerable rays converge and diverge. By describing the rays of light passing through each point, a complete description of the structure and distribution of light in that space is attained. The term *light field* was first introduced by Gershun in 1936,[22] but it is clear that Leonardo da Vinci understood light this way, when he wrote:

> The body of the air is full of an infinite number of radiant pyramids caused by the objects located in it. These pyramids intersect and interweave without interfering with each other during the independent passage throughout the air in which they are infused. (da Vinci, 1989, p. 50)

Such an understanding of light leads directly to, or perhaps arises directly from, the pinhole camera, or *camera obscura*. By filtering all but the rays of light that pass through a single point in space (the pinhole), the camera obscura reveals an image, and in doing so demonstrates that luminous space is infused with images. As da Vinci put it:

> The semblance of a body is carried by them as a whole into all parts of the air, and each smallest part receives into itself the image that has been caused. Immediately the air is illuminated it is filled with an infinite number of images which are caused by various bodies and colours located within it. (da Vinci, 1989, p. 50)

The convergent rays arriving at any point in the field contain a complete spherical image of the world as seen from that point of view. All around us is not empty air, but a dense field of innumerable spherical images. And all that we will ever see is a portion of the two spherical images centred on our eyes.

A light field can be formulated mathematically as a seven-dimensional scalar function. For each point in space (x, y, z) and for each direction (θ, ϕ), and for each wavelength λ, and time t, the light field L gives the intensity of light, i:

$$L(x, y, z, \theta, \phi, \lambda, t) = i \qquad \textit{Eq. 3 The Plenoptic Function}$$

[22] Translated by P. Moon and G. Timoshenko in (Gershun, 1939)

This is the *plenoptic function,* the formulation of the light field introduced by Adelson and Bergen (1991) in their work on the fundamental elements of vision. A physically complete (classical) description of a light field would require two further dimensions to capture the polarisation and phase of the light, but as these have relatively little importance in human vision, they are omitted here from the equation. Adelson and Bergen introduced the plenoptic function so that they might identify the elemental structures of light that lend themselves to vision. They show that, at the lowest level, vision can be reduced to measurement of local change along any of the dimensions of the plenoptic function. By cataloguing the various ways all the partial derivatives of this function might change, they arrive at a "periodic table" of elemental "visual substances" upon which vision is based.

However, the usefulness of the plenoptic function extends beyond that of its initial application. *All acts of vision, image capture, rendering, and image display can be construed as acts of sampling, recording, re-presenting or synthesizing the plenoptic function*. To see is to sample the plenoptic function. To capture a photograph, or record a film, is to sample and then save to some medium, the plenoptic function. To render a computer image is to sample a synthetic plenoptic function. And the purpose of an immersive visual display is to re-present the plenoptic function. As such, the plenoptic function provides a powerful framework for reasoning about immersion in visual media.[23]

[23] The concept of the light field has found application in a sub-field of computer graphics generally referred to as *image-based rendering*. In more traditional computer graphics, a scene is modelled as geometric shapes and surfaces, from which an image is synthesised by simulating, in some manner, the passage of light. With image-based techniques, there is no geometrical representation of the scene. Rather the scene is modelled as an array of light, similar to the plenoptic function, but typically with fewer dimensions in order that computations remain tractable. New images are synthesised by sampling this virtual plenoptic function. This class of techniques have come to be called "image-based" because the arrays of light are typically created by interpolating between many 2D photos of the world, which is to say that the underlying model is based not on geometry, but on images. See McMillan and Bishop (1995), Levoy and Hanrahan (1996), Gortler et al. (1996) and C. Zhang and Chen (2004). Attempts at photographically capturing the light field begin with Gabriel Lippmann's pioneering work on 'integral photography' in 1908 (Lippmann, 1908), while the first practical and commercially available light field cameras - the *Raytrix RX* (*www.raytrix.de*) and the *Lytro* (*www.lytro.com*) - would not arrive for another 100 years.

2.4.2 Vision and the Act of Perception

> The world is made of three-dimensional objects, but these objects do not communicate their properties directly to an observer. Rather, the objects fill the space around them with the pattern of light rays that constitutes the plenoptic function, and the observer takes samples from this function. The plenoptic function serves as the sole communication link between the physical objects and their corresponding retinal images. It is the intermediary between the world and the eye. (Adelson & Bergen, 1991)

The first fundamental insight provided by describing vision as an act of sampling the plenoptic function is that ordinary vision is an act of detecting structures and patterns in *all* dimensions of the plenoptic function, as opposed to the 'pictorial dimensions' of θ, ϕ and λ.[24]

The second fundamental insight is that at any single moment in time, our eyes sample a small subspace of the surrounding light field. By rotating our eyes we sample more of θ and ϕ, and we take in more of our surrounding. By shifting our head, or moving our body, we move in x, y, z, and see the world from different points-of-view. We are constantly shifting our coordinates within the volume of light within which we are immersed. *This is our normal mode of being in the world*. The parameters over which we sample the plenoptic function are, quite simply, never static, and at least a certain amount of our ability to make sense of our world is dependent on continuous *ego-motion* - our motion with respect to the light field.

This picture of vision matches that championed by J. J. Gibson in his ecological approach to visual perception (Gibson, 1979). In place of the terms "light field" or "plenoptic function," Gibson spoke of an "ambient optic array"; "the complete set of all convergence points, [constituting] the permanent possibilities of vision, that is, the set of all points where a mobile individual might be" (1979). "Air, in other words, is filled

[24] As this is of pivotal importance to the arguments presented in this thesis, the reader is encouraged to read Adelson and Bergen (1991) for a detailed discussion of how local change in all permutations of dimensions of the plenoptic function manifest as visual features of the environment.

with geometrical projections," wrote Gibson, and the visual system "is a system for *sampling* the ambient array" (Gibson, 1970, p. 75).

A *Gibsonian* notion of vision is predicated on perpetual movement through one's environment. We see the world *not from a point, but from a path*. He distinguishes between ambient vision, movements in θ and ϕ only, and ambulatory vision, which includes movements in x, y, z. But note that mathematically pure ambient vision - changing θ and ϕ without translating in x, y or z whatsoever - is very difficult in practice. When we turn our head, at least one of our eyes is moving in space. And even when we look about without moving our head, our eyes do not turn in their sockets exactly about their optical centres, but slightly shift in location. We are continually, both consciously and subconsciously, changing the domain over which we sample the plenoptic function.

> Perception is more than a means of passively representing the intrinsic physical organization of objects. Perception is inherently active and exploratory. It seeks out alterations in the vast flow of information enveloping it. These alterations are detected when the perceiver moves through the environment and probes it with a pair of glancing eyes. (Braund, 2008, p. 124)

This is the foundation of *perception-as-action*. We perceive the world by navigating the plenoptic function, and the trajectory we take is influenced by the values of the function itself. Movement begets perception which begets movement, and so on. This is not to say, however, that this moving and perceiving unfold sequentially, one after the other, but rather they unfold continuously and simultaneously. But even this is not accurate, for it suggests two simultaneous yet distinct processes. Rather, to move in one's environment, to travel a path through the plenoptic function, *is* perception.

> A perceiver's use of the ambient optic array is *active*, not *passive*. Accordingly, the eye is not simply a receptacle for discrete light stimuli imposed on the retina, but rather, an organ for exploring an ambient optic array. In this respect, the visual system is responsible for *obtaining* stimulation over time, not merely for *receiving* it [...] (Braund, 2008, p. 137)

This inseparable coupling of action and perception, to the degree that it is nonsensical to talk of one without the other, is most elegantly demonstrated in the c*entrifugal governor,* a device that simultaneously senses and influences its environment.

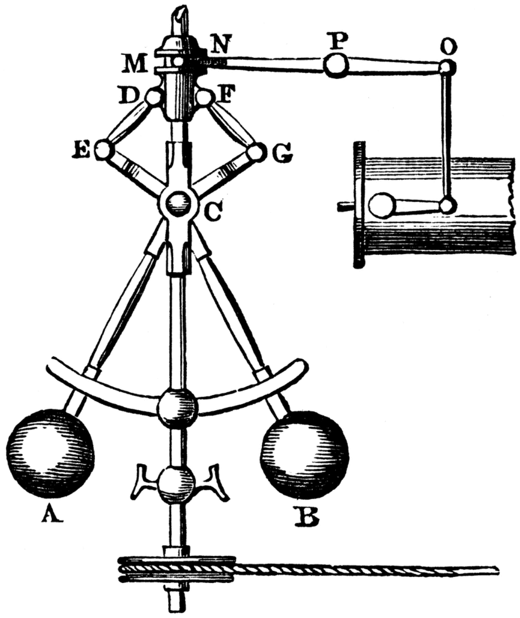


Figure 2 - The Centrifugal Governor, illustrated in Whitney (1895, p. 2585).

A centrifugal governor is often used to govern the speed of an engine. As the engine gains speed, centrifugal forces fling out the two arms (A and B), which push the lever (O) to slow the engine, which slows the rotation, which lowers the arms, which pulls the lever to speed the engine, which raises the arms.... and so on. However, this is not a sequence of events, as the sequence of words makes it seem, but a continuous equilibrium of forces. There is no taking of turns between sensing (moving the arms) and acting (moving the arms), for they are one and the same. To change speed *is to* measure speed, and vice versa, and it is simply unintelligible to speak of one independently of the other. Like this, the governor demonstrates how measurement and control – perception and movement – can be inseparably unified.

> The boundary between perception and action frequently fades: many actions are undertaken for their perceptual consequences, and perception is often tuned to those aspects of the world that are available for the observer to act on. (Wexler & Van Boxtel, 2005)

Perception-as-action is central to Gibson's philosophy of perception, but is not uniquely his. It is also central to cybernetics and perceptual control theory (where actions are not seen as effecting change in the world, but as effecting change in perception, i.e. "*behavior is the control of perception*" (Powers, 1973). It has found voice in notions of the "embodied mind" from Varela (e.g. Varela et al., 1991), and is central to the work of Alva Noë (e.g. 2004) and other active sensorimotor theories of perception (e.g. Mossio & Taraborelli, 2008). It is also advocated by theorists of *dynamic cognition*, for whom the centrifugal governor has been adopted as the quintessential metaphor of cognition (e.g. Van Gelder, 1995).

2.4.3 Presence as active perception of the world

It is the view of the author that this manner of being-in-the-world presented here is an essential element of presence. Undoubtedly our being-in-the-world is a complex affair and this is by no means the complete picture, but if we are to be present in the virtual world, then it is this aspect-of-being - this physical relationship with the plenoptic function – that must be replicated.

Here, presence is not identified as a feeling or self-perception. Being present, after all, is our normal mode of being in the world, and as such, largely invisible to us. It is certainly available to introspection, such that "if you stop and think about it you will (no doubt) perceive that you are physically present in some environment'' (Schloerb, 1995, p. 65),[25] but whether or not we are consciously reflecting on or feeling presence is incidental to the relationship our perceptive system is engaged in with the environment. This is really just a matter of definition, for one could equally define presence as the feeling of perceiving the world this way. This additional level of indirection seems however to be unnecessary, and somewhat at odds with the direct approach to perception stalwartly advocated by Gibson.[26]

---

[25] Quoted in Zahorik and Jenison (1998).

[26] On the topic of self-awareness, Gibson wrote: "perception is an awareness of the world. An awareness of the self *accompanies* it but does not *contribute* to it." (Gibson, 1970, p. 79)

Defined as such, presence can be regarded as a constant unwavering feature of being in the real world. It grows neither stronger nor weaker with the varying emotions or beliefs of the viewer. This is not to deny emotions or beliefs a role in presence entirely, for there are perfectly valid definitions of presence in which they do play a role, but to clearly demarcate this particular picture of presence as independent of these particular qualities of mind.

Unlike other definitions, presence here is not predicated on the construction of some inner mental model of the world. This is a recurrent theme in presence theory, that spatial presence results somehow from a process of constructing or acquiring a mental representation of the world as in, for example, Biocca et al. (2001) or Schubert (2009). In the definition suggested here, no such inner representation of the world is required. That is, presence is independent of acquisition of knowledge about the world, and whether or not we emerge with a coherent or correct inner representation of the world is considered irrelevant. This picture of presence is compatible with Gibson's insistence that perception is *direct.* Gibson eschews completely suggestions that perceptions arise from inference or interpretation of stimuli, or that there is a difference between sensations and perceptions, or that all we experience is an inner representation of the world, or that even our experience with the real world is somehow mediated. Rather, perception is a direct access to and awareness of the world (e.g. Gibson, 1972, 1979). Or as a dynamic cognitivist might say, perception is a coupling of two dynamic systems; just as the centrifugal governor is coupled with an engine, the dynamic system that is the mind is strongly coupled with the dynamic system that is the world.[27]

2.4.4 Presence and Binding

At this point, the reader may protest that while ego-motion is certainly frequent, it is not completely without interruption. Sometimes, after all, we are still. The reader may even be trying this experiment right now; head held perfectly still, their eyes fixed on some distant, stationary object, juxtaposed with a very close foreground object as a

[27] The example given by Gibson (1966) is that of a simple analogue radio receiver. The inner state of the radio is strongly coupled (tuned) with its environment (radio waves).

measure of stillness. With some concentration, perhaps they will succeed staying the shift in parallax between the foreground and background, and for a brief time hold a stationary position in the plenoptic function. But take note not just of how difficult this is to achieve, but of the physical sensations that accompany this.

"Vision is always a mixture of proprio- and extero-ception" (Reed, 1988, p. 290). Our physical movements do not just influence which parts of the plenoptic function we see, but are a source of information in themselves. Physical actions trigger somatosensory[28] signals that are integrated with visual signals to form unitary percepts (e.g. Gepshtein & Banks, 2003; Sambo & Forster, 2009; Reuschel et al., 2010). That is, vision depends on more than the light striking the retina, but also on complex "extra-retinal" signals.[29] For example, proprioceptive signals generated by muscles in the eyes as they focus and converge on an object are in themselves depth cues (Wexler & Van Boxtel, 2005). We quite literally see distance with these muscles.

Further, these extra-retinal cues not only augment visual cues, but have been shown to alter how we perceive visual stimuli. Carefully conducted experiments where both a moving and stationary observer are subjected to identical visual stimuli have shown that the moving observer, despite experiencing the same visual stimulus as his stationary counterpart, perceives the same three dimensional structure differently (Wexler, Lamouret, et al., 2001; Wexler, Panerai, et al., 2001; Wexler & Droulez, 2003; Wexler & Van Boxtel, 2005). The "same optic flow can lead to very different perceptions of 3D shape when generated by the observer's own movement than when generated by object motion" (Wexler & Van Boxtel, 2005).

How we perceive the incoming light is influenced by the state of our body, and if indeed it is possible to view the world from a fixed location in the plenoptic function,

---

[28] The somatosensory system includes perception of touch, pressure, pain, temperature, position, movement and vibration.

[29] "There are two main sources of extra-retinal action-related signals in the brain. One consists of a copy of the motor command, known as efference copy or corollary discharge, and exists only in the case of actively generated motion. The other consists of reafferent feedback signals from the vestibular organ, and somatosensory or proprioceptive signals from the muscles. These neural extra-retinal signals related to action are combined with visual signals in multiple brain areas" (Wexler & Van Boxtel, 2005).

that is, to see an *image*, it would be accompanied by some very strong physical extra-retinal signals.

This leads to a second aspect-of-being considered vital for presence: the fusing, or *binding*, of multiple stimuli into single unitary percepts, and the subsequent binding of percepts into a single, coherent and unified experience of being in the world. At every moment of everyday experience, a multitude of sensory cues, or 'perceptual objects,' arising from distinct proprioceptive and exteroceptive sensory mechanisms, such as texture and colour, size and shape, position and orientation, shading and shadows, reflections and occlusions, sounds and reverberations, motion and ego-motion, and so on, are bound into a single, seamless experience of the world. This "multitude of different kinds of perceptual objects" are perceived ultimately as a "single perceptual and behavioral space" (Revonsuo, 1999). In cognitive science, the question of how this binding might occur, how diverse cues arising from within a single sense modality and across the senses are integrated into single, unitary conscious experience, is known as the *binding problem.* The binding problem often arises in the search for neural correlates of consciousness, where the challenge is to explain how "temporally and spatially segregated activity in neuronal ensembles is reassembled in order to generate a seamless conscious experience" (Harvey & Sanchez-Vives, 2005).

This successful binding of sensory inputs is taken here, along with active-perception, to be an aspect of being in the world that is essential to presence. The goal of immersion after all is to stimulate the viewer with an assortment of conceits - colours, shapes, sounds - that are ultimately perceived not as a patchwork or cacophony of stimuli, but as a coherent world. It would seem then that the goal of mimetic immersion is to present stimuli that are amenable to this process of binding. Or as Harvey and Sanchez-Vives (2005) put it, for presence to survive, "the constellation of sensorial cues in a virtual environment must be in accord with some basic rules which, in the real world, govern the relationship between sensory events." That is, we must concern ourselves not just with the sensorial cues, but with the relationship between cues.

An understanding of these 'basic rules' - the conditions under which binding occurs, or does not occur - can be directly employed in the creation of immersive experiences. For example, two binding phenomena that are routinely employed in immersive systems are stereoscopy, where two images are perceived as one, and stereo sound, where two sounds are perceived also as one. In both cases, the conditions under which fusion does or does not take place have been very carefully charted, and this knowledge is directly applicable to the design of immersive experiences.

What other binding phenomena might play a role in presence? Like presence itself, the term 'binding' is a polyseme for many different mental or neurological phenomena. Many of these different forms of binding have been distinguished by observing their absence in neurologically impaired or injured patients. Pathological failures of binding emerge as different forms of *agnosia* - highly specific forms of selective blindness in which the sufferer is unable to see or recognise specific aspects of the world, such as motion, shape, objects, colour, despite being able to perceive the individual percepts that make up these phenomena. Apperceptive agnosia, achromatopsia, prosopagnosia, semantic dementia, akinetopsia and simultanagnosia or Balint's syndrome, for example, can all be considered failures of binding (Revonsuo, 1999). Binding mechanisms include:

- *Spatial grouping.* The grouping of visual stimuli into structures. This form of binding is captured in the Gestalt laws of perceptual organization (e.g., proximity, similarity, closure, symmetry, common fate, continuity) (Revonsuo, 1999).
- *Property binding (feature integration)*. The binding of distinct elementary features within a single sensory modality. For example, the colour, texture, depth, shading and shape cues arising from an object are integrated to produce the perception of single, coherent object (Treisman, 1996).
- *Part binding.* Multiple parts of an object are recognised as belonging to the same object (Treisman, 1996).
- *Multi-sensor binding.* The binding of stimuli between different sensors of the same sense modality. For example, between two eyes or between two ears.

- *Multi-modal binding.* Binding between the different sense modalities. For example, touch, proprioception and vision might all combine when we discern the shape and weight of an object held in hand (Biocca et al., 2001).
- *Serial binding (event binding).* Ability to recognise the continuity of percepts over time (Revonsuo, 1999), or the binding of simultaneous stimuli as arising from the same event. This form of binding may play a role in the perception of events, and so be implicated in causal perception - the perception of cause and effect (see for example Cavazza et al. (2007)).
- *Semantic-conceptual binding (cognitive binding).* Recognition of objects. The binding of percepts to semantic knowledge or memory of the world (Revonsuo, 1999).
- *Location binding.* The binding of spatial perceptions with all of the above (Treisman, 1996).

Binding should not be confused with the traditional psychological notion of *association*. While they may both be described as subconscious and automatic, an important distinction between the two is that once binding takes place, access to the individual percepts is completely surrendered. For example, once stereo fusion takes place, the two individual images are entirely extinguished, and cannot be reclaimed without breaking the bind.[30] In contrast, if seeing a red raincoat triggers memory of a certain film,[31] the red raincoat does not suddenly become imperceptible. (For more on the problem of association and direct perception, see Chapter 13 of Gibson (1966)).

> In sum, phenomenal experience is unified. It is unified globally, to form one coherent phenomenal whole (the unity of consciousness), and it is unified locally, so that objects are experienced as coherent sets of phenomenal features located in specific spatial locations. The different kinds of binding and disintegration at the phenomenal level suggest that normally the contents of

---

[30] Stereo fusion serves here as an example, because it is one of the minority of bindings that occasionally fail. Most bindings, for most people, are perpetual, and the constituent signals are simply never perceived.

[31] The author is incapable of seeing a red raincoat without thinking of Nicolas Roeg's *Don't Look Now*.

consciousness are the result of a great variety of binding mechanisms that, to some extent at least, function independently of each other. (Revonsuo, 1999, p. 180)

The position held here is that failures of binding due to conflicting or incoherent sensory cues are impediments to presence. The art of immersion then lies in the presentation of conflict-free intra- and inter-sensory cues. An obvious failure of binding, for example, is when stereo fusion fails, and the world is revealed to be two flat images. Or its sonic counterpart, when summing localisation fails and we hear two sounds from distinct locations in space rather than one, this too can be considered a failure of binding.

For immersive displays however, it is the binding of proprioceptive and exteroceptive cues that are held to be particularly important to preservation of presence, for it is these types of bindings that are most frequently disrupted or prohibited by limitations in visual display. In the real world, a natural and inescapable consequence of active perception is that all movements (including staying stationary) give rise to a continual binding of proprioceptive and exteroceptive cues, and it is through the union of these cues that we perceive the world. As will be seen below, the capacity of a display to provide coherent proprioceptive and exteroceptive cues is a measure of its *immersiveness,* and that it is this very binding, or the lack of it, that is the greatest impediment to presence with contemporary display technology.

Cross-sensory interactions

Biocca et al. (2001), who also draw a connection between perceptual binding and presence, focus on two particular aspects of binding between of inter-sensory cues: *cross-modal enhancement* and *cross-modal transfer*.

Cross-modal enhancement occurs when "stimuli from one sensory channel enhances or alters the perceptual interpretation of stimulation from another sensory channel. [This] might include changes in detectability, perceived intensity, perceived fidelity, or some other perceptual quality of the stimuli from another sensory channel" (Biocca et al., 2001). For example, concordant visual stimulus has been shown to greatly influence the perceived location of sounds (see Section 3.10.9). Similarly, and perhaps more surprisingly, is that the inverse seems to be true; for a sudden sound that provokes an involuntary change of visual attention, such as turning the head, has been shown to improve subsequent perception of visual stimuli in the locale of the disturbance (McDonald et al., 2003). Cross-modal enhancement between the aural and visual domains is used to great effect in the AVIE and *La Dispersion Du Fils*.

Cross-modal transfer occurs when "stimulation in one sensory channel leads to the illusion of stimulation in another sensory channel" (Biocca et al., 2001). Cross-modal transfer can be considered, therefore, a form of synaesthesia. Examples include perception of inertia and weight when manipulating a virtual object, despite lack of any haptic stimuli, or hearing two objects collide, when in fact they are silent. In the real world, synaesthesia is a rare occurrence and largely considered abnormal. However, in the virtual world, the viewer is frequently subjected to contradictory or incomplete stimuli of the like that would simply never occur in the real world. The latter case, where sensory cues are missing or poor, invites a "filling in" of the missing sensations in a manner that occurs perhaps only rarely in everyday life. As a result, this form of synaesthetic transfer can be considered particular to mediated experiences.

Through a series of experiments, Biocca et al. (2001) demonstrate that users experiencing a greater amount of spatial presence are more likely to experience cross-modal phenomena in virtual environments. Whether it is presence that depends somehow on these cross-modal phenomena, or vice-versa, or both arise from some other underlying mechanism is unclear. Nonetheless, the phenomena of cross-modal enhancement and synaesthesia are documented in immersive virtual environments, and present not only avenues for understanding presence, but potentially powerful tools for the creation of immersive experiences.

Here, these two cross-modal interactions - enhancement and transfer - are considered to play important roles in the fabrication of immersive experiences. At the very least, they provide mechanisms for masking flaws in a mediation, allowing the viewers perceptive system to fill in the blanks or connect the dots, so to speak. The key, therefore, lies in avoiding conflicts that prohibit such cross-modal transfers.

2.4.5 Invariants and the Structure of Vision

Gibson's ecological approach to perception describes a relationship with the plenoptic function, and this relationship is taken to be an essential component of presence. But what does it tell us about the structure of the plenoptic function itself? Can Gibson's approach to perception provide insight into those aspects of the real world that must be faithfully replicated for presence to be preserved? In other words, does Gibson's theory provide insight into the algorithms - the physics - by which the virtual world is rendered into physical stimuli, as described in the previous section?

"A perceptual system does not respond to stimuli, [...] but extracts invariants" (Gibson, 1976, p. 236). As we move within the plenoptic function, (and we are always moving), or objects or sources of light move in the world about us, the pattern of light impinging on the eyes changes. This change is orderly, in that there are rules governing the manner in which the structure of the light changes, and these rules, which are determined by our biology, our manner of being-in-the world and the physical nature of the world itself, are fixed. These rules manifest as *invariants* - relationships and ratios, transformations and deformations, accretions and deletions that describe constancy in both the structure (structural invariants) and evolution (transformational invariants) of the plenoptic function.[32] Vision is "seeing the non-change underlying the change" (Gibson, 1971, p. 32). Now we begin to see the true importance of action and movement in perception: only with variance can there be invariance.

> Action is a necessary requirement to obtain perceptually relevant information, and no perceptual ability can occur if invariants specified by action are not available. (Mossio & Taraborelli, 2008)

Persistence and change detected in the plenoptic function by a moving observer simultaneously provides information about the position and path of the observer in the world (perspective structure) and the world itself (invariant structure). Gibson saw that

---

[32] In his three major works *Perception of the Visual World* (1950), *The Senses Considered as Perceptual Systems* (1966) and *The Ecological Approach to Visual Perception* (1979), Gibson provides over two dozen different examples of invariants. For a very concise summary of these, the reader is referred to Goldstein (1981).

> optical change over time in the optic array is what specifies both the locomotion of the observer and the motion of an object whereas the *non-change* over time in the changing array is what specifies the spatial arrangement of the environment and of the object in the environment. (Gibson, 1970)

Or as fellow ecological psychologist E. S. Reed put it:

> Perspective structures specify where we are heading, and invariant structures specify the nature of what we are heading toward. (Reed, 1988)

A fundamental task of vision, therefore, is to distinguish between perspective structures - transformational invariants arising from our motion through the world - and change arising from motion in the world itself. For example, how does the perceptive system discern between true motion parallax arising from ego-motion, and identical visual transformations that might be caused (by chance or design) by moving objects rather than a moving observer? The answer has already been given above: the binding of somatosensory and exteroceptive cues. Invariants can be both intra-modal (relating to one sense only), but most importantly, they can be multi-modal, capturing regularly occurring patterns or structures in data fused from all or any of the senses. Perception, therefore, is the detection of invariant structures across all modalities and between *efferent, reafferent* and *exafferent* sensory patterns:

> perceptual systems are able to discriminate between reafference (sensory input resulting from self-motion) and exafference (sensory input produced by external events) in virtue of their relation to efference (internal information elicited by self-motion). Perceptual systems receive at the same time reafference and efferent copies generated by a given movement and such information is used to perceive a specific action as self-initiated. (Mossio & Taraborelli, 2008)

We are able to distinguish between self-induced motion cues and a moving world because, once somatic cues are included, the multi-modal invariant structures arising in the former are simply *nothing like* those arising in the latter. To confuse the sight of moving through the world with a moving world is like confusing an orange and apple – it is only possible if you look to their shape and ignore all the other cues. Perspective structures are reafferent – they are not just detected by our actions, they are *created* by our actions, and therefore instantly distinguishable from exafferent perceptions.

2.4.6 Ecological Optics

So according to an ecological approach to perception, we perceive the world by detecting invariant structures, and these structures are both modal (within one sense modality) and multi-modal (between sensory modalities). The proposition made here is that in order for presence to be preserved in a virtual environment, it is these structures that must be preserved by the mediating technologies. These structures are considered invariant because they adhere to laws, and it is because of their nomological invariance that they are seized upon by the perceptive system. Gibson termed these laws *ecological physics*, with constituent branches of optics, acoustics and haptics, and took great pains to distinguish them from classical physics.

> Ecological optics does not have to be concerned with the problem of waves or particles nor with the laws of refraction, reflection, and diffraction. It is primarily concerned with margins, borders, contrasts, ratios, differences, and textures in the array. (Gibson, 1961)

Gibson's optics structures are, at the lowest level, composed from those elemental "visual substances" described by Adelson and Bergen (1991), namely recognisable patterns of local change in the plenoptic function. As such, a key difference between physical optics and ecological optics is frame of reference, for the former is largely allocentric and concerned with structures and phenomena in the world that exist independent of an observer, while ecological optics is entirely egocentric, and concerned only with the structure and evolution of the plenoptic function of an observer.

Now, in the previous section, it was argued that mimetic immersion demanded a nomological, processual prescription of how the virtual world manifests as physical stimuli, and how these stimuli evolve over time. It was suggested that this was tantamount to specifying a virtual physics of light, sound and force (i.e. optics, acoustics and haptics) for the fictional world. And it was supposed that if these processes were a faithful emulation of real-world physics, then presence would be sure to be preserved in the virtual world. This was the brute-force approach to immersion labelled physical mimesis. The presence hypothesis, however, allows for the possibility that a different physics be employed, provided it encapsulates those features of real-world physics that are salient to presence. The proposition put forward here is that these features are exactly those prescribed by Gibson's ecological optics - the laws describing the structure and evolution of variants and invariants as an observer moves through the optical array.

Zahorik and Jenison, who also draw on Gibson's ecological approach to perception, arrive at a similar conclusion:

> Preservation of similar invariant structure for the additional sensory modalities seems to be at least a starting point in attempting to characterize natural perception/action coupling. Perhaps support of these invariants in a virtual or remote environment might represent a set of minimal conditions for the perception of lawful environmental response to action. (Zahorik & Jenison, 1998, p. 88)

Ecological optics will be discussed further in the context of *La Dispersion Du Fils,* for it will come to play a role when the question of immersion in fictional, abstract worlds is addressed.

2.4.7 Summary

Active perception is the act of creating and then detecting invariant patterns and structures in multi-modal stimuli. This manner of perceiving the world *constitutes* presence. That is to say, presence does not arise from an active-perceptive relationship with one's environment, it *is* this relationship. Although this discussion has focused on vision, it is for no other reason than explicatory simplicity, and that the panorama is primarily a visual medium. When we describe perception as the act of creating and then detecting regular patterns and structures in multi-modal stimuli, it is a description that applies equally to all sensory modalities.

Presence is a continuous unwavering feature of being conscious in the real world. As this is our normal mode of being in the world, it is inescapable, persistent and largely invisible. Presence becomes obvious when this method of perceiving the world is disrupted. In an immersive mediation, there are a number of possible disruptions; limitations on our freedoms to act or explore the environment, missing or contradictory cues arising from movement in the environment, which result in failures of binding, or flaws or defects in the mediation that result in a breaking of the laws of 'ecological physics.'

Further, the twin mechanisms of cross-modal enhancement and transfer have been identified as features of perception that may be exploited in the design and creation of immersive experiences. The key is to avoid perceptual conflicts that may prohibit these cross-modal mechanisms.

2.5 IMAGES AND IMMERSION

Armed with a firmer definition of presence, it is now possible to approach the subject of immersive displays and interfaces.[33] In particular, the challenge of immersing multiple viewers within the same shared space is examined. This discussion begins by identifying the requisite qualities a display must possess if it is to give rise to presence. These criteria are described in terms of the plenoptic function, and it is shown that the *immersiveness* of a display can be directly linked to the *degrees of freedom of vision* that it supports. While a true light field display would permit the exercising of all these degrees of freedom, thereby ensuring preservation of presence, such a display is beyond current know-how.

In lieu, the *image* must be adopted as a surrogate for the light field, and it is from this surrogacy that many of the obstacles to immersion arise; for the perception of the world offered by an image is utterly unlike that enjoyed in the real world. Nonetheless,

[33] Again, the focus will be on the perception of light, although the arguments that follow are equally applicable to the perception of sound. Indeed, one need only exchange the light-field for the wave-field to arrive at almost identical conclusions for immersive sonic display.

it is possible to construct an image-based immersive display by understanding and then mitigating the impact of the image on vision. Such solutions, however, tend to be suitable for a single viewer only, and it would seem that a conflict between immersion and multiple viewers is at hand.

The *stereo panorama* is introduced as representing the ideal compromise between the two. By sacrificing certain degrees of freedom of vision, the stereo panorama provides a semi-immersive experience for multiple viewers.

The perceptual experience of the stereo panorama is examined, along with its strengths and shortcomings as a medium of the light-field. Methods for alleviating these shortcomings are then discussed, including the simulation of ego-motion and interactivity. Finally, a warning on the propensity for interaction to destroy presence, rather than enhance it, is given.

2.5.1 An Immersive Display

What are the necessary or sufficient qualities that a visual interface must possess in order that a viewer may enjoy a sense of presence in a virtual environment? Previously, presence was defined as an active-perception of one's environment, in which the viewer engages in the continual creation, detection and subsequent binding of invariant structures and patterns in the surrounding multi-sensory field of information. An immersive display, therefore, can be defined as a display that permits this same perceptive relationship with a virtual environment. This definition has a number of components:

- ❖ For the plenoptic function to remain representative of the virtual world, the display must preserve the ecological structures which are important to vision. Variants must remain variant, and invariants must remain invariant.
- ❖ To perceive the light field is to perceive structures in *all* dimensions of the plenoptic function. As structures and patterns in some dimensions of the plenoptic function are only revealed, or only exist, when moving through the light field, ego-motion is an essential element of presence.

- Multi-sensory stimuli must give way to a coherent, unified and conflict-free percepts and experience. These stimuli must be amenable to binding, which is to say that the constellation of reafferent (external cues resulting from action), efferent (internal information elicited by action) and exafferent cues (external cues produced by the external world) must observe the same essential relationships exhibited in the real world.
- In particular, the physical actions and motions of the viewer should give rise to conflict free binding of efferent cues and reafferent cues, in such a manner that perspective structures are not only created by action, but *intentionally* created and *anticipated* by the viewer.

In Section 2.5.6 below a further component will be added to this list, which we presage now for completeness:

- A display must permit the perception of the virtual world as a *perceptual rest-frame*.

An interface possessing all these properties could then be rightly described as immersive. This definition does not preclude the possibility of a display possessing only some of these features, or possessing them in degrees, in which case it is possible to talk of displays with differing levels of immersiveness.

Light Field Display

A naive approach based on *physical mimesis* would seek to replicate the light field in its entirety. In doing this, all the criteria of presence would be satisfied, as the virtual environment would be visually identical to its real world counterpart.[34] It would be holography in its purest form, with the passage through the air of each and every one of the innumerable rays of light faithfully replicated, unbounded in space and infinitely detailed for every possible moment in time.

[34] Recalling the notation introduced previously, we might write that the light field L_V produced by $M(V_E)$ is identical to the light field L_E of the real environment E. i.e. $L_V = L_E$

A more tractable approach is that introduced in the previous section as *metameric mimesis*.[35] As the artificial light-field need only be *perceptually* indistinguishable from the hypothetical original, it need only match the spatial, spectral, temporal domains of the human visual system.[36] For example, a sequence of still images presented in fast enough succession will be perceived as continuous motion, a feature of human vision upon which television and cinema are founded. Hence the plenoptic function need not be uniquely defined for all t, but at some suitably regular intervals (e.g. 60 or 120 times per second). Similarly, the limited spatial acuity of the eye allows a reduction in resolution of θ and ϕ, without any perceptible loss of information.

For the spectral domain, a simplification is offered by way of colour metamers. For a trichromat,[37] light composed of just three suitably chosen wavelengths (λ_R, λ_G and λ_B) is perceptually indistinguishable from light composed of an arbitrary spectrum of wavelengths. This allows us to rewrite the plenoptic function:

$$L(x, y, z, \theta, \phi, t) = \langle R, G, B \rangle$$

where R, G, B are intensities of fixed wavelengths λ_R, λ_G and λ_B

Eq. 4 The Trichromatic Plenoptic Function

In doing this, the plenoptic function is reduced from a seven dimensional scalar field to a six dimensional vector field. Further the resolution and range of R, G and B need only match the colour and intensity acuity and dynamic range of the human eye.[38] Further reductions in complexity are attained by putting bounds on the domain over

---

35 i.e. L_V != L_E, but S`=S.

36 This should perhaps be rewritten with different emphasis, for the capabilities of the human eye still far outreach that of our best display technologies. The *challenge* is to create an artificial light-field that approaches the resolution and range of the spatial, spectral, temporal domains of the human visual system.

37 Almost all humans are trichromats. Many birds are tetrachromats. Intriguingly, evidence that a small percentage of women may be tetrachromats has recently come to light (Jacobs et al., 2007).

38 Which is no mean feat. The dynamic contrast ratio of human vision is extraordinarily broad, ranging from 10^{-6} to 10^6 cd/m^2. However, this is achieved through dynamic control of sensitivity - constriction or dilation of the pupil, a shift from photopic to scotopic vision, and narrowing the eyes to see in bright light - and so the range of brightness perceivable within a single moment is not as pronounced. This is another feature of human vision that may be exploited by a metameric display - using dynamic exposure to simulate high dynamic range.

which the plenoptic function is defined. For example, the range of x, y, z over which the light field is reproduced can be reduced to just the space over which the viewers are likely to roam.

An ideal immersive display is therefore reasonably well defined. We need only replicate the light field with a resolution and range commensurate with human vision and human motion. Such a display could be described as both pervasive and passive, for the light field exists everywhere (within a certain volume of space), and is independent of the position and orientation of the viewer. It would, therefore, be implicitly compatible with multiple simultaneous viewers. Theoretically, such a light field display would be flawless, and with no flaws to reveal it, invisible.

Unfortunately, such a flawless light field display remains an ideal. The spatial, spectral and temporal resolutions and range of contemporary holographic displays are far below those of vision (Yaras et al., 2010).

Display Artefacts

When a display is flawed, it distorts the light field. These distortions, which are sometimes referred to as *display artefacts*, can be concisely described with plenoptic function. The most common digital display artefacts arise from insufficient range or quantisation of the spatial (x, y, z), angular (θ, ϕ), spectral (λ or <R, G, B>) and temporal (t) dimensions of the plenoptic function. For example, pixels can be understood as an insufficiently grained discretisation of θ, ϕ, while the restriction of θ, ϕ to a range narrower than that of human vision manifests as a *frame*. It is possible to construct a table:

Display Artefact	**Display Feature**	**Limitation**
Frame	Field of view	limited θ, ϕ
Pixels	Resolution	discrete θ, ϕ
False colours	Colour gamut	limited λ or <R, G, B>
Colour/Intensity Banding	Colour depth (bits per pixel)	discrete i or λ or <R, G, B>
Low contrast or brightness	Dynamic range	limited i or <R, G, B>
Noise	Signal to noise ratio	random error in i or <R, G, B>
Lag	Latency	offset in t
Flicker, jitter	Frame rate	discrete t

Table 1 - Visual Display Artefacts

These are just some of the ways that an imperfect display may corrupt the signal, and in doing so, reveal itself.

Quantisation, Boundaries and Invariants

What are the minimum permissible resolution and range for each of the dimensions of the plenoptic function? If presence were reliant on the apparent "disappearance of the medium," then these permissible resolution and ranges would be tied to the visibility of display artefacts, and therefore a function of human visual acuity. This is not the position held here. Rather, it is proposed here that the minimum permissible resolution and range for each of the dimensions of the plenoptic function is tied to Gibson's concepts of variant and invariant structures. The structures of the plenoptic function that are important to vision (thresholds and gradient flows, deletions and accretions, spatial ratios and relationships), must somehow be preserved and, most importantly for presence, must be preserved when the user undertakes motion through the plenoptic field. If these features depend on calculation of local rate of change (and change of rate of change, and so on), it can be seen how over-discretisation and aliasing of the function can potentially destroy these features, or create unwanted features.

The most important implication of this is that the ranges and resolutions required to preserve these variant and invariant structures may be well within the range of visibility. It may be that aberrations in the light field, such as pixels, or colour banding, or flickering can all be perfectly visible, and yet have no, or little, diminishing effect on presence. That is, it is not the visibility of these artefacts that poses a threat to presence, but their perturbing and destructive influence on Gibson's invariant structures. This is a more concrete restating of the *presence hypothesis*: that the mediated light field may significantly differ from the original, and yet presence is preserved.

Now, as vision is predicated on measurement of local change in *all* the dimensions of the plenoptic function, insufficiently fine-grained quantisation of *any* of these dimensions is detrimental to vision. However, not all dimensions of the plenoptic function are sampled in the same manner. It may be that presence is more sensitive to limitations or resolutions imposed on some dimensions than others. For example, we know from black and white imagery[39] that when λ is discretised to just two values, many images remain perfectly intelligible. On the other hand, if angular resolution were reduced to just two values the result would be akin to an image of just 2x2 pixels, which is clearly insufficient for vision as we know it.

Sampling theory (e.g. the Nyquist limit) dictates that any concept of a 'minimum resolution' will be tied to the size or frequency of structure within the light field. That is, any lower limit on resolution will be dependent on such things as the scale and shape of the virtual world, the speed of motion (including the viewer) and even the size of the viewer relative to the virtual structures. As such, generic statements about the minimum resolution are not made easily. However, the restrictions imposed on the various dimensions by current display technology tend to be of an *all-or-nothing* nature, with some dimensions enjoying resolutions of hundreds or thousands of gradations, while others are confined to just one or two samples.[40] So while all contemporary practical display technologies impose limitations on the resolution and boundaries of the plenoptic function, it is those limitations that *completely extinguish* a) visual structure, or b) the ability to discover new structures, that have the most pernicious effect on presence. As such, the question of sufficient resolution is entirely supplanted by the more fundamental question of *degrees of freedom* and *dimensions of vision*.

The 10 Degrees of Freedom of Binocular Vision

Our ability to discover new structures in the plenoptic function is dependent on the manner with which we shift our coordinates within the function. At this point a clarification is needed on the "dimensionality" of vision. At any moment in time, our position in the plenoptic function is not defined by a single coordinate (x, y, z, θ, ϕ), but by a set of points. Each of the eye's photo-receptors are stimulated by a bundle of light-rays which occupy a small slither of the plenoptic function, and it is the union of all these small subspaces that defines the *visible set* of the plenoptic function at any one

[39] i.e. bi-chromatic, not greyscale.

[40] A perspective image presented on a typical present-day computer monitor quantizes θ, ϕ by 1920 x 1080, $<R, G, B>$ to 256^3 discrete values, and time t by 60Hz. The spatial dimensions x, y, z however are restricted to a single point in space or, in the case of a stereo image, two points in space.

time.[41] The shape of this visible set is determined by many parameters, some fixed and some variable over time. For example, the size and position of the photo-receptors within the retina can be considered fixed. So too the distance between the eyes.

The geometric *variables* of a binocular observer are 10 in number: 6 variables describe the position (x, y, z) and orientation of the head (θ, ϕ, ψ), 3 variables define a common point upon which both eyes converge (vergence) *and* focus (accommodation)[42] (f, u, v) and therefore define the direction of the eyes, and 1 variable defines the size of the aperture of both the eyes (a).[43]

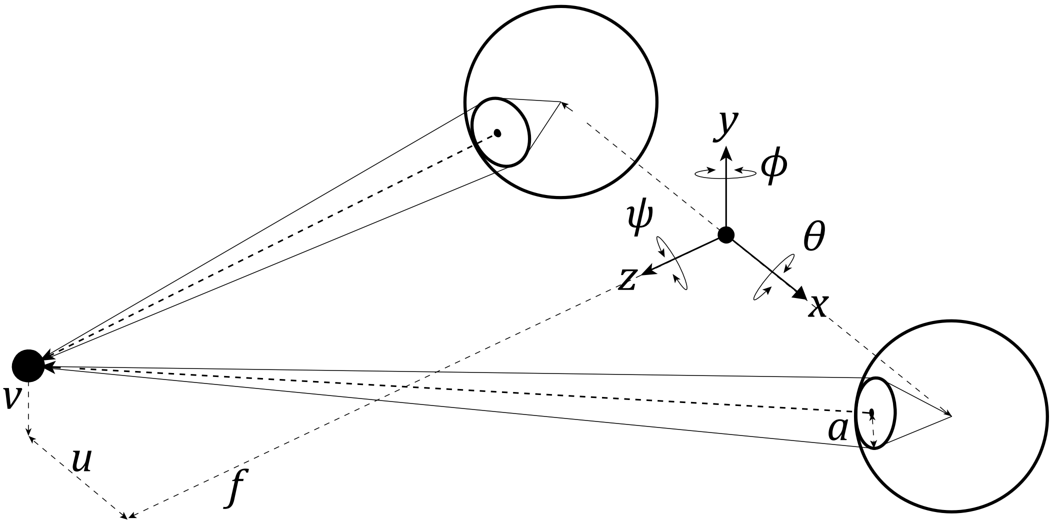


Figure 3 - The 10 degrees of freedom of binocular vision.

These are the variable *degrees of freedom* with which an observer actively explores the plenoptic function. They define the visible set: the set of coordinates used to sample the plenoptic function during vision at any one instant. As the viewer has 10 degrees of freedom, and the plenoptic function only 5 spatial dimensions, there is a non-trivial mapping between the degrees of freedom of vision and dimensions of the plenoptic function. Some enjoy a one-to-one mapping: translating the head in x, y or z correspondingly translates the visible set in x, y or z. Likewise, pitching the head up or

[41] Note that the region over which the plenoptic function is sampled is not necessarily continuous (there are gaps between photoreceptors, as well as the retinal blind spot), nor uniformly dense (higher density at the fovea than the periphery).

[42] See Banks et al. (2012) for more on vergence and accommodation.

[43] For simplicity it is assumed that both eyes share the same pupil size. A more complete model might also include the possibility of squinting, but this is omitted here.

down (θ) translates the visible set in θ. Turning the head left to right (ϕ) translates the visible set in ϕ, but it also slightly alters the position of the eyes as they rotate around a central axis, and so also influences the x and z dimensions of the visible set.

The tilt of the head ψ, however, is not so simple. First note that, if one assumes the retina to be completely circular, then single eye rolling about its optical axis has no effect on the visible set, and no new information or structures within the plenoptic function can be revealed by such an action. It is true that, because of the occluding effect of the brow and nose, the field of view of a single eye cannot be considered circular, and for the binocular viewer, roll of ψ is always accompanied by a shift in position of at least one of the eyes, so tilting the head will actually produce a shift in all of the x, y, z, θ, ϕ coordinates of the visible set. However, by and large, such an action tends to reveal little new information about the world. The principal effect of rolling the head lies rather in the perception and understanding of familiar objects such as faces or text.[44] This can be seen by observing use of head roll in normal vision, where it is seldom exercised, except when we wish to inspect an upturned object in a more familiar orientation.

As for focal length and aperture, it is important to note that the human eye is not a pin-hole camera, but admits light through the pupil, the diameter of which changes with available light.[45] As such, normal vision is not a sampling of the plenoptic function from a single x, y, z coordinate, but over a small circular region of x, y, z, the size of which is determined by the diameter of the pupil. The pupil's non-point-like diameter is the cause of depth-of-field, so restricting the visible set to a single x, y, z coordinate effectively transforms the eye into a pin-hole camera with infinite depth of field, thereby removing this element of depth perception from vision.

Changing eye direction (u and v) has largely the same effect as changing head direction (θ and ϕ), although there are some very subtle differences. When we look askew without turning the head (u), the distance between the eyes (the stereo baseline)

---

[44] This is known in psychophysics as the *inversion effect*, in which detection, recognition or understanding of certain visual phenomena is greatly impaired when they are viewed upside down. It is pronounced in perception of faces (Farah et al., 1995) and biological motion (Troje & Westhoff, 2006).

[45] A typical pupil diameter ranges from 4mm when constricted to 9mm when fully dilated.

is effectively diminished, while when we turn our head and keep our eyes looking forward, the interpupillary distance remains constant. Further, as we move the eyes about while holding the head still, the field of view undergoes change as the brow and nose come into view.

With respect to presence, it cannot be presumed that all 10 degrees of freedom are equal. In this respect, the most important aspect of a particular degree of freedom is whether it can be intentionally exercised to reveal more of the plenoptic function. Vergence, accommodation and dilation and constriction of the pupil are involuntary dimensions of vision that cannot be controlled directly, for vergence and accommodation are automatic responses to foveation, while aperture is an automatic response to the amount incoming light. Further, the altering of vergence, accommodation and aperture has relatively minor effects on the visible set, as does the tilting the head. If the degrees of freedom are divided into those that have a major or minor impact on the visible set, the ambulatory (x, y, z) and panoramic (θ, ϕ) degrees might be considered major, while ψ, u, v, f and a could be described as minor.

An ideal display would allow the viewer the freedom to vary all or any of these 10 variables at will, providing them with a visual experience identical that of vision in the real world. With a non-ideal display, there will be some degrees of freedom that are met with incorrect stimuli, which includes the common cases that the stimuli remains constant regardless of the viewer's motion, or no stimuli is provided at all. Immersive displays can, therefore, be characterised by the number of degrees of freedom for which they provide correct stimuli. These 10 degrees of freedom of vision, which correspond to Slater's "sensorimotor contingencies" or "valid sensorimotor actions" (Slater, 2009), provide a metric by which different visual displays might be compared, or classified in equivalency classes according to the degrees of freedom they offer the viewer.

2.5.2 Pictorial Perspectives

Unfortunately, true light field displays remain an ideal. In lieu, we must make do with a more ancient technology: *the image*. Rather than viewing the light field directly, an image is recruited as an intermediary. Perception of the light field is now a two-step process: the light-field is projected onto a surface to form an image, and the viewer samples this image as if it were the light field itself. In this section the consequences of such an arrangement with respect to immersion and presence are examined.

An image, in its most general sense, is a mapping of the plenoptic function onto a two-dimensional surface.[46] When this mapping can be described as the intersection of rays with a surface, it can be called a projection, and when this image surface is a plane, it is a planar projection. A *single-point* projection is one where all rays pass through a single focal point, or *centre of projection*. This is classical single-point perspective, the pin-hole camera, the most familiar of all projections.

When the viewer is exactly coincident with the centre of projection, it would seem that an image can act as a faithful surrogate for the light field. This is a false impression.

The first effect of the image is the collapsing of all samples to a single point x, y, z. With this collapse, all information in these dimensions is extinguished. This includes the instantaneous spatial cues such as stereo depth and monocular depth of field, as well as temporal cues, such as motion parallax. The image eliminates completely the possibility of using rates of change in any of the spatial dimensions of the plenoptic function to derive structure from motion. Second, as the light field has been projected onto a surface, the eye must now focus and converge on this image surface and nowhere else, so focal length and vergence are also eliminated from the act of vision.

All that remains is the information in the pictorial dimensions of θ and ϕ. But if the image is smaller than our field of vision, such that the entire image may be consumed in a single glance (which is almost always the case), our ability to turn our heads to reveal more of the plenoptic function vanishes completely. *The last degrees of freedom have been extinguished.* All the information available within the image is now instantaneously

[46] There are, mathematically speaking, any number of ways such a mapping might be realised. For a summary of the properties of different mappings (or 'cameras' as they are known) the reader is referred to Yu et al. (2010) and Neumann et al. (2004). Seitz and Kim (2002) (and to a certain degree Pajdla (2002)) provide a complete analysis of stereo mappings - images in which horizontal binocular disparity is preserved - and show that the family of all possible stereo mappings is limited to projections where rays of light lie on planes, hyperboloids, or hyperbolic paraboloids.

available, and no matter how much you shift and wriggle in the light of an image, no new information will be revealed.[47] At this point, perception ceases to be an action, and our relationship to the light field becomes purely passive. There is no longer any role for the body in the act of perception; it has been excised from the equation, and we have become disembodied passengers.

More than this, however, is that whenever we move, or exercise any of our degrees of freedom of vision, *the only visual structures revealed are those that betray the image as a surface*. Looking up, down, left or right not only draws attention to the border of the image, but reveals the real world outside of this image. Shifting position, rather than revealing the spatial structure of the virtual world, now only serves to further reveal the geometry of the image surface, and destroy any perception of structure that might exist. By adopting the image as surrogate, ego-motion has been transformed from an act of revelation to an act of disillusion. This is also true of binocular disparity, vergence and accommodation, three perceptive faculties for perceiving depth in the world, all of which now only serve to reveal the image as a surface.

This is the vision of the world afforded by the image. It should hopefully be clear that it is nothing like that of ordinary vision. This is not to say that the image is completely ineffective as a mediation of the light field. After all, this is the manner with which the light field is presented in photographs, television and cinema, which are all evidence of our ability to recognise and understand the world represented in an image. However, it seems that understanding or recognising the world represented in an image has little to do with presence, and it is this that concerns us here.

So, given the crippling effect that the image as a medium of the light field has on immersion, how might an image be used to construct an immersive display? Can any of the artefacts of the image be diminished? Can any of the dimensions of vision be somehow reinstated in the act of perception?

---

[47] It is true that we constantly move the eyes about as we foveate on various elements of an image. However, when the image is encompassed completely by our field of view, foveation does not reveal new regions of the plenoptic function.

Brunelleschi's Peepshow

Methods for reducing the visibility of the image surface are as old as linear perspective itself. Filippo Brunelleschi (1377–1446), having painted two outdoor scenes with the increasingly popular method of linear perspective, invited his audience to behold these images in a very specific manner. Antonio di Tuccio Manetti, Brunelleschi's biographer, recounts how Brunelleschi

> had made a hole in the panel on which there was this painting; . . . which hole was as small as a lentil on the painting side of the panel, and on the back it opened pyramidally, like a woman's straw hat, to the size of a ducat or a little more. And he wished the eye to be placed at the back, where it was large, by whoever had it to see, with the one hand bringing it close to the eye, and with the other holding a mirror opposite, so that there the painting came to be reflected back; . . . which on being seen, . . . it seemed as if the real thing was seen: I have had the painting in my hand and have seen it many times in these days, so I can give testimony. (Translated by J. White, 1987, p. 116).

A similar method of viewing images would be used two centuries later by Dutch painters such as Johannes Vermeer and Samuel van Hoogstraten in the construction of "perspective cabinets," small wooden boxes containing painted scenes that are only to be viewed with one eye through a small peephole (Balzer, 1998). It is also the function of the 18th century *Zograscope*, a lensed device through which a viewer would view with one eye specially designed prints (Chaldecott, 1953).

The reason for such elaborate contrivances is that viewing an image through a peephole is known to reintroduce stereopsis – the perception of depth. When an image is viewed through a peephole, the illusion of depth is qualitatively similar to the illusion of depth provided by binocular stereo, if perhaps less powerful in magnitude (Vishwanath & Hibbard, 2013). The mechanisms behind it are complex, and still open to debate, but a simplified explanation is that viewing through a simple peephole removes much of the perceptual evidence of the image surface, thereby allowing the depth cues embedded within the image itself to dictate the perception depth.[48]

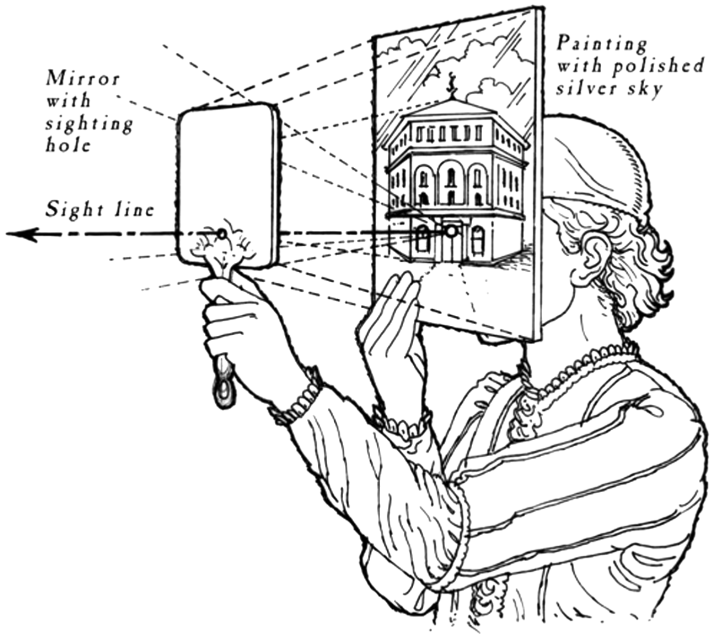


Figure 4 - Brunelleschi's Peephole (source unknown).

The simple peephole achieves this by influencing vision of the image in a surprising number of ways; for there are no less than five separate mechanisms at work. First, the peephole constrains the viewer to the correct centre-of-projection, so the image is guaranteed to be a valid geometrical representation of the light field. This removes any distortions in the image that might betray the image surface. Second, in curtailing all possible movement of the viewer's point of view, it conceals the fact that the image lacks motion parallax and contains no information in the x, y, z dimensions of the plenoptic function. Third, in permitting strictly monocular viewing, the binocular cues of disparity and vergence no longer play a role in vision, and so can no longer play their part in revealing the depth of the image surface. Fourth, it occludes the sharp edges of the image, which are often the cues that provide the most precise judgment of the distance and shape of the screen surface. Fifth, it creates the illusion that the scene extends indefinitely behind and beyond the edges of the peephole, while

[48] A more nuanced explanation involves the interplay between the relative and absolute depth cues. See Vishwanath and Hibbard (2013).

simultaneously removing the real-world from view. This allows the image to be perceived as the *visual background*, a concept that will be discussed further below.[49]

Brunelleschi's peephole provides important lessons for the design of immersive displays. It makes evident the pernicious effect that the image surface and its border have on perception of the scene contained within the image, and demonstrates concretely how these effects can be defeated.

Stereo Images

When not one, but two images are used to mediate the light field, binocular depth perception - the perception of depth afforded by the simultaneous perception of the world from two disparate points of view - is reintroduced into the act of vision. This has two consequences. First, in providing binocular disparity for the scene, the strongest depth cues of the image surface are eliminated. The depth cues within the scene now outweigh those of the image surface, and the conflict of depth cues is resolved in favour of those within the scene itself. The result is not just that the scene is seen with depth, but that the image surface is rendered invisible in the process. As will be seen below, it is this disappearance of the image, an oft underappreciated feature of stereo imagery, that is perhaps most important to presence.

Second, it reintroduces a single, minor degree of freedom into the act of vision: vergence. As the viewer trains their view on different elements of the stereoscopic image, their eyes automatically converge or diverge to match the binocular disparity of the elements under scrutiny.

However, the addition of stereo to media such as cinema and television has a relatively minor impact on presence, for a number of reasons. Vergence is an involuntary reaction to foveation, and altering vergence does not reveal new regions of

[49] A sixth mechanism may also be at work. The mirror of Brunelleschi's original experiment, and the lens of the Zograscope both have the effect of allowing the viewer to focus at a greater distance than the apparent distance to the image plane. Kubovy (1988) suggests this may even be the case for a simple peephole without mirror or lens, for the small peephole may act as an aperture-stop, and thereby increase depth of field, allowing the viewer to focus further away than the image plane and still perceive the image in sharp focus.

the plenoptic function. As such, vergence, in of itself, is not a degree of freedom of vision that can be intentionally exercised to explore ones visual environment. Secondly, the real world, which remains clearly visible all about the image, firmly remains the dominant rest-frame. As such, stereo images, while enriching the act of perception, still remain a passive mode of perception. It is only when stereo imagery is coupled with other immersive techniques that it becomes a more important feature of the illusion.

Vergence-Accommodation Conflict

It is also necessary to note an inherent flaw in stereo images, which is the unnatural decoupling of vergence and accommodation. While the eyes converge at a distance dictated by the binocular disparity in the stereo image, which may be anywhere in front of or far beyond the screen, the eyes must maintain their focus on the image plane. There is now a conflict of depth cues, as both accommodation and vergence are providing incoherent extra-retinal depth information. This vergence-accommodation conflict, as it is known, has been shown to cause discomfort and eye fatigue and to diminish the perception of depth (Kooi & Toet, 2004; Hoffman et al., 2008; Ukai & Howarth, 2008; Shibata et al., 2011; Banks et al., 2012). Shibata et al. (2011) calculate a "zone of comfort" for viewing stereo images, where the range of acceptable vergence distances is given as a function of focal distance. For a screen at 5m, for example, it is suggested that the virtual object not be brought closer than approximately 1.3m.

In the listing of the 10 degrees of freedom of vision above, vergence and accommodation were compounded into a single degree of freedom to reflect their coupling in normal vision. A stereo image, therefore, can be said to provide the viewer only one half of this degree of freedom.

First-person Perspectives

If the position and orientation of the viewer's head or eyes can be accurately ascertained, and this data can be used to retrieve or create the correct projections of the light field for the viewer, then an image can act as a faithful surrogate of the light field, even for an active observer (Kubitz & Poppelbaum, 1973) (Fisher, 1981). For this to be viable however, the images must be calculated or retrieved with such speed that the delay between measurement of the viewer's point of view and presentation of the correct image is imperceptible. It is this rapidity of image formation that explains why, despite the algorithms for forming images from light fields being thoroughly understood since the 14th century,[50] *immersive* images would not appear for another 600 years, when a sufficiently rapid means of rendering these images would become available in the form of the computer. In other words, with regards to the use of images as a conduit of the light field, the fundamental breakthrough offered by the computer is not so much that it can synthesise images, but that it can do so in real-time.

This is the fundamental premise behind the head-mounted display (Heilig, 1960; Sutherland, 1968) and immersive screen-based systems such as the CAVE VR system (Cruz-Neira et al., 1992; Cruz-Neira et al., 1993) and its numerous derivatives. In the former, two image displays are physically attached to the viewer's head, ideally in a manner that they completely fill the viewer's field of vision, while in the latter, a seamless mosaic of image surfaces surrounds the viewer. In both cases, the viewer's coordinates in the plenoptic function are tracked in real-time, and the correct imagery is presented for their point of view.

Both classes of systems can potentially provide correct stimuli for 8 ½ of the 10 possible degrees of freedom of vision. The missing ½ degree is due to vergence-accommodation conflict, for regardless of the depth in the virtual scene, the eyes must always remain focused on the image plane. The other missing degree is that of aperture (a). It is relatively simple to render images with an artificial depth-of-field and dynamic exposure that correctly reflect changes in aperture, however this demands some method of tracking not just the viewer's head, but also the direction of their gaze

[50] Early practitioners of perspective were predominately interested in an algorithmic understanding of their art, where the process of image creation could be reduced to a series of rules. It is for this reason that da Vinci's writings on subject of perspective and painting read very much like an introduction to computer graphics.

and size of the pupil. While technically possible, in practice this is seldom done, and so a typical CAVE or HMD offers just 8 ½ degrees of freedom.[51]

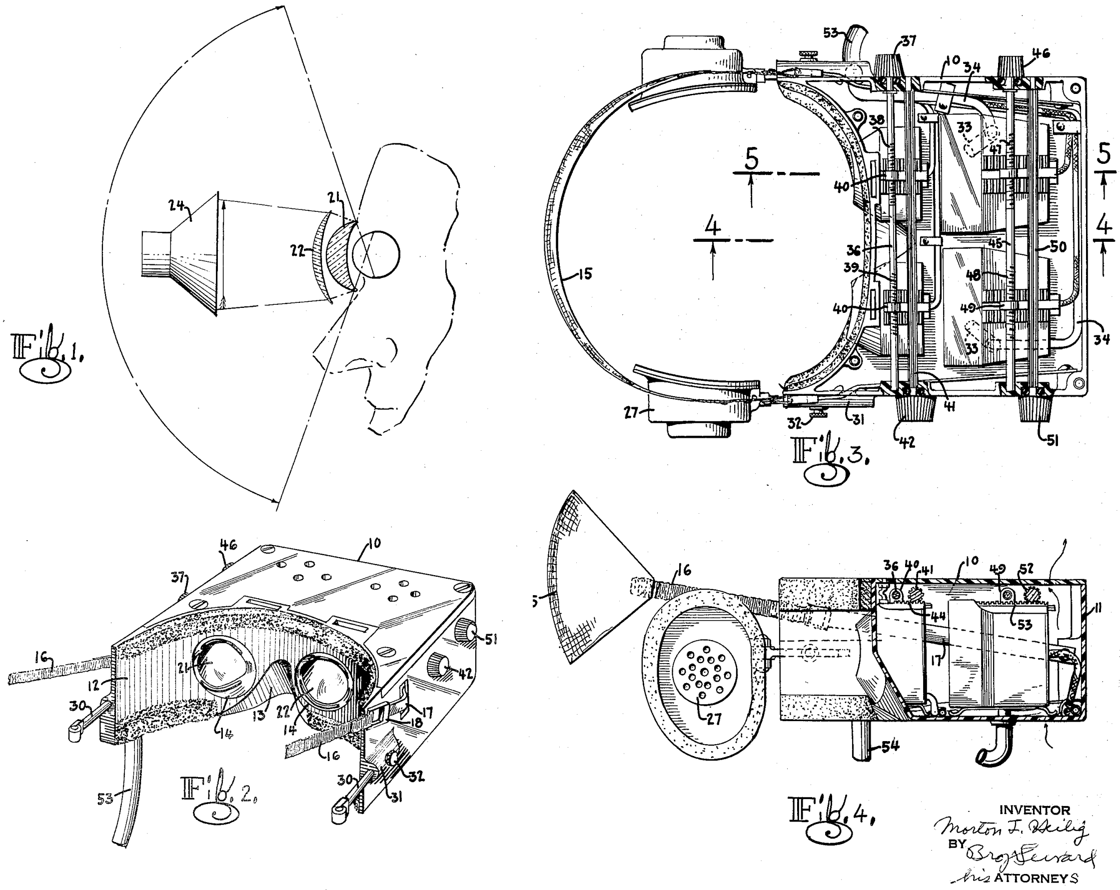


Figure 5 - Morton Heilig's Telesphere Mask (1960)
An example of a head-mounted display in which two image displays - in this case miniature televisions - are physically attached to the viewer's head. Here they are viewed through wide-angle optics, providing a 140 degree field of view.

This highlights the direct mapping between the capabilities of the tracking system and the vision of the world provided the viewer. A device that can only measure the orientation of the viewer's gaze, for example, can only provide the viewer with the freedom to move in θ and ϕ of the plenoptic function. In contrast, a device that can measure the position of the viewer head or eye's can provide the viewer access to the x, y, z dimensions of the plenoptic function. As such, the development of first-person

[51] New generation of HMD may begin to incorporate eye-tracking, allowing not only the simulation of depth-of-field and exposure, but also the exploitation of the non-uniform acuity of the human eye to reduce rendering complexity by lowering resolution in the periphery of vision.

immersive systems is intimately tied to the problem of tracking, and the history of virtual reality is as much the history of tracking solutions as it is of image display.

The most important difference between the head mounted display and screen-based systems such as the CAVE lies in the different manner in which one perceives one's self within the virtual world. With screen-based systems, the viewer perceives their own body naturally and directly. However, a basic limitation of such systems is that the viewer's body is inescapably superposed on top of the imagery at all times, meaning the regions of the light field that lie between the body and the eyes can never be displayed. In contrast, in the head-mounted display the viewer is deprived of a body altogether, and must be provided with a virtual surrogate.

The fundamental limitation shared by both of these methods is that they are limited to a single viewpoint.

For the head-mounted display, the obvious remedy is to provide each viewer with their own device. The challenge then lies in providing all users with virtual representations of bodies that permit natural communication and interaction between one another. Such natural communication demands that the physical appearance, movements and gesture, eye movements and facial expressions of all the viewers be faithfully replicated within the virtual world, and although piecemeal techniques for tracking these features are emerging, to date only a crude representation of the user's physical body within the virtual world has been possible.

For screen-based systems, support for multiple viewers can be found in the "multiplexing" of multiple images on the same screen (Fröhlich et al., 2005; Pross et al., 2012). The technical challenge faced here is effectively that faced by any stereoscopic display; the delivery of unique images to different eyes. This can be seen by noting that any stereo screen can equally be considered a generic 2-channel display and, rather than provide a single stereo point-of-view for one viewer, it could present two unique monoscopic points-of-view to two viewers.[52] By using a combination of various

[52] Sometime in the 1950s, Du Mont released the *Duoscope* television which used polarised light to allow viewing of two different television channels on the same screen at the same time (Horzu, 1954).

stereoscopic channel separation methods at the same time (polarity, wavelength, temporal – see Section 3.6), it is possible to present multiple stereo images on a single screen. In what can be considered the state-of-the-art of this approach, Kulik et al. (2011) demonstrated the display of six unique stereo points of view on a single screen. However, their system required one projector for each viewer, so if the technique were used to build a multi-user 6-sided CAVE, 6 projectors would be required for every viewer, which might be considered prohibitive. The same ratio of 6 projectors per viewer would also be met, were one to adopt this technique with the AVIE.

Image multiplexing is a promising technique, but for the moment resists extension to large numbers of viewers. Is there an alternative? Could perhaps a certain degree of immersion be sacrificed in exchange for a simpler, more extensible approach to multiple viewers?

2.5.3 Panoramic Images

A panoramic image is formed by projecting the light field onto a cylindrical surface. Such an image offers a view of the light field with an unbounded horizontal field of view. The immersant may now explore freely and without limitation the ϕ dimension of the plenoptic function. The panorama, by eliminating the bounds of the horizontal field of view, injects a single, vital degree of freedom into equation of vision. In doing so, perception is reinstated as *act*; an act that is both physical and intentional. Perception becomes, once again, a volitional and physical activity.

The difference between zero and one degree of freedom is nothing less than embodied perception.

Further, when constructed with a sufficiently high vertical field of view, all boundaries of the image are largely beyond view, thereby removing one of the strongest clues that the image is a surface. This is one of the reasons that Robert Barker would go to great lengths to conceal the upper and lower edges of his painted panoramas, by constructing viewing platforms with canopies and fences that naturally occluded the boundaries of the image. In fact, it is precisely this feature of his invention that receives the most attention in the text of his patent:

> There must be an inclosure within the said circular building or framing, which shall prevent an observer going too near the drawing or painting, so as it may, from all parts it can be viewed, have its proper effect. This inclosure may represent a room, or platform, or any other situation, and may be any form thought most convenient, but the circular form is particularly recommended. Of whatever extent this inside inclosure may be, there must be over it (supported from the bottom, or suspended from the top,) a shade or roof, which, in all directions, should project so far beyond this inclosure, as to prevent an observer seeing above the drawing or painting, when looking up; and there must be without this inclosure another interception, to represent a wall, paling, or other interception, as the natural objects represented, or fancy, may direct, so as effectually to prevent the observer from seeing below the bottom of the painting or drawing, by means of which interception nothing can be seen on the outer circle, but the drawing or painting intended to represent nature. (Barker, 1796).

2.5.4 Stereo Panoramic Images

Nevertheless, despite the tell-tale boundaries of the image lying beyond view, a panoramic image is still perceived as an image, for binocular disparity continues to betray it as a surface. Perhaps, if a *stereo panorama* could be constructed, then the remaining perceptual evidence of the image surface would vanish, and make way for the depth cues within the scene itself. This is the promise of the stereo panorama.

Can stereo and panorama be combined? Traditional stereoscopic images are easily created by capturing the world from two points of view, each displaced to the left and the right of one another according to some direction of view. Were we to capture two single-point panoramas, one displaced to the left of the other, and then display the pair as a stereo image, the stereo effect would only be correct for one direction, non-existent for other directions and completely inverted in another. How then might one create a stereoscopic panorama?

The problem stems from the assumption of a singular direction of view, the very idea of which is anathema to the panorama. The solution lies therefore in abandoning such a notion, and adopting a direction of view that varies across the image surface. If we assume the direction of view to be everywhere perpendicular to the screen surface, and then displace the left and right viewpoints laterally from this viewing direction, a stereo panorama is attained. The consequence, however, is that the image rays no longer pass through a single centre-of-projection, but rather the centre-of-projection is now defined by a circle. The stereo panorama belongs to the more exotic family of images known as multi-perspective or multiple-centre-of-projection images, images whose points-of-views are, in fact, not points at all.[53]

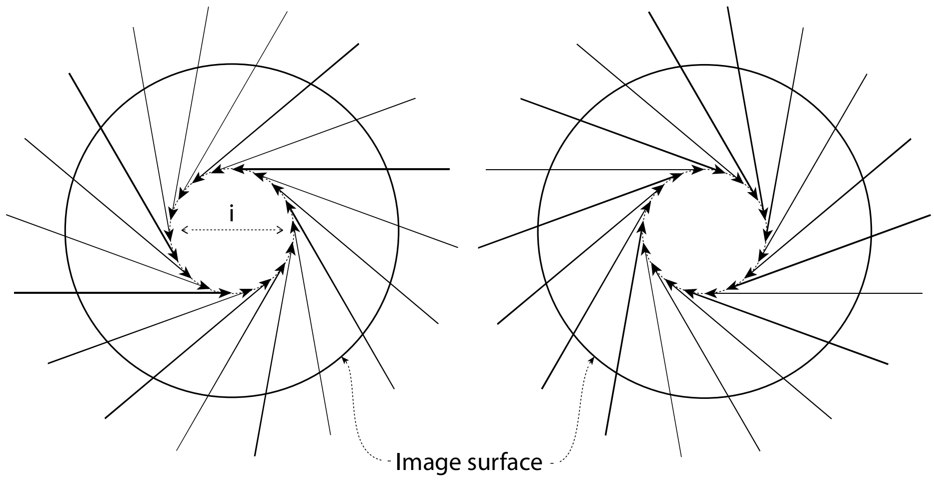


Figure 6 - Overhead view of the omnistereo projection for the left and right eyes. i - the radius of the circle of projection - determines the stereo baseline, and is typically set to 6.5cm to match the average distance between eyes.

This particular form of image projection has come to be known as *omnistereo*, and it is the method of display employed by the AVIE. The creation and display of omnistereoscopic images will be discussed in greater detail in Section 3.9, but for the moment it is useful to jump ahead and describe a little the *experience* of the stereo panorama. The following discussion is grounded on the personal experiences and observations of the author, as well as the reported experience of hundreds of visitors to the AVIE. It is not a reflection of any one AVIE application, but rather draws on observations of all iCinema AVIE experiments conducted over the years. Note however that no attempts to rigorously or quantitatively measure aspects of the user's experience of the AVIE were ever conducted, for while this would certainly be a worthwhile

[53] Any image projection that captures the plenoptic function over a region of x,y,z, rather than a single point, is considered to have multiple centres-of-projection.

undertaking, it has always remained beyond the remit of both this thesis and the research program of the iCinema centre as a whole.

When omnistereoscopic imagery is displayed in the AVIE, the walls of the cinema quite tangibly disappear. A vista opens up, extending in all directions and the sensation of space stretching out beyond the cylinder is both physical and concrete. As predicted, the combination of stereoscopy and large field of view results in the disappearance of the image surface. It is not so much that the screen becomes a window, for this would suggest the presence of a window frame, but rather that the screen, and the whole AVIE apparatus with it, tend to disappear. In fact, of the real world, all that remains is two strange black disks, one beneath the feet and the other floating mysteriously in the air. The virtual world is not just imagined to be complete and encompassing, extending above and below the floating circles of the floor and the ceiling, but it is *directly perceived* as such.

2.5.5 Distorted Visions

The panorama provides an effectively undistorted view of the light field for a viewer standing at the centre of the cylinder.[54] For the non-central viewer, however, the panorama no longer serves as a geometrically correct projective surrogate of the light field. The result is an apparent distortion of space. As the viewer moves about, space around them appears to compress, dilate and shear, and as the viewer moves away from the centre of projection, these distortions grow in magnitude.

It is natural to anticipate that such a pervasive distortion of vision might be highly detrimental to any sense of presence. In practice, however, this is not the case, and there are a number of reasons for this.

First, note the continuous nature of the deformations. In contrast to the deformations experienced in multi-planar projections such as the CAVE, where the

[54] Effectively, but not exactly, for the omnistereo projection produces incorrect binocular disparity in the periphery of vision. However, it has been shown by Couture et al. (2010) that these peripheral errors are below the stereo acuity and stereo field-of-view of normal vision, and therefore can be largely ignored. In addition, the 3D glasses worn by the viewer tend to obscure the periphery of vision, further concealing any errors.

corners and edges of the image surface play havoc with the perceived geometry,[55] here there are no abrupt discontinuities whatsoever. The importance of this stems from the definition of an immersive display presented above (Section 2.5.1), in which the introduction of invariants was considered detrimental to the preservation of presence. With a discontinuous image surface, false invariants such as corners, and angles, edges and junctions – the elemental building blocks of vision – are injected into the visual scene. In such cases, it could be said that the *topology* of the scene is not preserved. The geometric errors caused by these surface discontinuities are particularly disruptive when the virtual world is in motion, or when the viewer is in motion, for their constancy amid change tends to dominate the perception of the scene, and instantly reveal the shape of the image surface.

Second, the distortions described above were based on a geometric interpretation of the image, as if we see by reprojecting the image back out into the world, from the viewer's point of view. They provide a *naive* notion of how the viewer might perceive the distorted space, based solely on the geometrical interpretation of stimuli. In practice, perception of these distortions and the subsequent experience of space are not governed by these projections, for the human visual system is furnished with a remarkable capacity to automatically correct distorted stimuli. To appreciate this, it is useful to return for a moment to the perception of images.

La Gournerie's Paradox

Previously, some time was taken to explain how the vision of the world offered through an image was in fact utterly unlike natural vision, so much so that it is reasonable to ask, if the view of the light field offered by the image is so alien, and so susceptible to distortion, why is the image such a ubiquitous and effective means of communication? There are some obvious reasons: an image can be presented on paper or in a book, for one, and reproduced mechanically. In addition, the single-point planar projection provides a means of image creation that is sufficiently simple that it can be

[55] All CAVE systems present an image that is correct for a single viewer only, while everyone else is subject to a distorted view. Yet, strangely, it is commonplace for CAVE demonstrations to be given to groups of people, in which case it is guaranteed that everyone, but one, sees nothing but a fractured vision of the world.

performed by hand, or automated with a simple optical device. There is, however, a more subtle reason for the pervasiveness of the perspective image.

When an image is used as a projective surrogate for the light field, it is assumed that it represents the view of the world from the viewers current point of view, and the space and structure represented in the image can be reconstructed by simply re-projecting the image from our point of view back out into the world. As such, one would expect that as a viewer moved away from the point of view captured in the image, the perception of space within the image - the shape, size, positions and orientations of objects - would all visibly distort. This is the perception of an image predicted by a geometric interpretation of vision (Woods et al., 1993; Held & Banks, 2008; Pollock et al., 2012).

Somewhat surprisingly, however, this is not the case. By virtue of some perceptive mechanism of human vision, a perspective image, when viewed from some point other than the centre of projection, *appears undistorted*. The viewer's perception of spatial structure within the image - the shapes and forms, ratios and angles, relative positions, dimensions and orientations and so on - tends to be completely unaffected by the discrepancy in viewpoints. More than this, space within the image exhibits a form of constancy, so that when the viewer actively moves about, the perceived space and structure within the image appears to remain invariant. Despite the geometry of vision prescribing otherwise, a viewer's perception of an image is not dependent on them adopting a specific point of view.

> Perhaps the most fortuitous attribute of picture perception, then, is that one may view a picture from many locations other than the point of composition, and distortions of virtual space will interfere little with the perception of the picture's content. If this were not true, the utility and appreciation of pictures would be vanishingly small. (Cutting, 1986b, pg. 552)

It is this feature of image perception that makes possible the reliable communication of 3D spatial structure with a 2D representation. It is also the feature that makes possible the use of the image as a *shared medium*. Shared in two senses: a) because the viewer need not be at the centre of projection to view the image, a multitude of viewers can partake of the same image at the same time, and b) because, regardless of their point of view, everyone sees more or less the same forms. This democracy of perspective underlies the entire enterprises of cinema and television, where a single image might be shared by a vast auditorium of viewers and they all more or less perceive the same projected forms. After all, recollections of a film or television program are seldom prefixed with the phrase "from where I was sitting …"

(This automatic and unconscious distortion correction could also be regarded as the final step in the disembodiment of the viewer. Not only do we have no influence on the perspective of the scene within the image, but we are even denied a role in distorting the image.)

The perceptual constancy of the perspective image has long been a source of intrigue, and even today the exact mechanisms underlying it remain mysterious. Michael Kubovy named it "the robustness of perspective" (Kubovy, 1988) while James E. Cutting, who wrote about this "viewpoint nonspecificity of pictures" extensively (Cutting, 1986b, 1986a, 1987, 1997), referred to it as *La Gournerie's Paradox*, after Jules de la Gournerie, who may have been the first to analyse the phenomenon mathematically (De La Gournerie, 1859; Pirenne, 1970).

There are a number of competing theories seeking to explain how and why the mind performs this trick. The interested reader may find an overview of the contending theories in Banks et al. (2005), in which the majority of theories can be classified in one of two classes: the *pictorial-compensation* and *surface-compensation* hypotheses. In the former, it is suggested that the observer is somehow able to rectify their perception of space by making use of information and structure within the depicted scene itself. For example, parallel lines in the world produce vanishing points in the image, and three such vanishing points are sufficient to reconstruct the image's centre of projection, which is then somehow used to rectify distortions (De La Gournerie, 1859; Adams, 1972; Greene, 1983; Kubovy, 1988).

In contrast, in the latter class of theories, this rectification takes place not by extracting information from the scene, but by perceiving the image-surface itself

(Gibson, 1950; Pirenne, 1970; Perkins, 1973; Rosinski & Farber, 1980). The evidence for the *surface-compensation* hypotheses is compelling. Banks et al. (2005); Vishwanath et al. (2005) demonstrate with a certain irrefutable rigour that our ability to correct image distortion and see the forms within the image as constant even as we move about, is dependent on our ability to judge the *local orientation* of the image surface. If, and only if, the local slant of the image surface can be perceived, will the scene within the image perceived undistorted. They suggest that judgement of image surface orientation is afforded first and foremost by binocular depth, and then by the geometry of the image border. Other clues that reveal the surface of the image are texture or pattern of the image surface, scratches or stains, picture grain or pixels, and uneven illumination, reflections or specular highlights.

The consequences of the surface-compensation hypothesis on our stereo panorama are not trivial. Undistorted perception of space depends on the visibility of the image as a surface. And yet, it was argued above, that to see depth within an image, the image surface must be imperceptible. Here lies one of the paradoxes of the immersive image. The exact cues we wish to eliminate in order to increase immersion are the very same cues we must preserve for an undistorted vision of space. We must choose, it seems, between an undistorted two-dimensional image, and a distorted three-dimensional space. Or as Kubovy put it: "We must perceive the window in order to see the world" (1988, p. 61).

Perception of Space in the Stereo Panorama

For a viewer moving about within the AVIE, space appears to distort. As the viewer moves about, the background and foreground visibly shift in a form of reverse motion parallax, and space itself appears to curve and flex. As they approach the screen, space compresses, or as they recede, it dilates.[56] Accuracy of spatial judgement is diminished, and without recognisable objects to provide a sense of scale or distance, it would be very difficult for a viewer to ascribe an accurate numerical distance between themselves and some distal object.

[56] See Woods et al. (1993) for a geometrical account of such distortions on a planar screen.

These are all indications that the screen-surface has disappeared and with it the robustness of perspective. For the moving observer, the sensation is strange indeed. The apparent motion of clearly stationary objects provokes a form of visual paradox, not unlike an optical illusion. The world tends to *swim,* but when the viewer focuses their attention on any particular object, no motion can be perceived. Most important, however, is that the perceived 3D space stubbornly resists collapsing back into an image.

For a stationary viewer, the experience of the AVIE is markedly different. For when the viewer stops moving, these spatial distortions tend to vanish. Certainly the accuracy of spatial judgement remains diminished, but the qualitative sense of space, of the relative size, distance and position of shapes and their intended straightness or curvature, all tend to appear undistorted and correct. Most importantly, even when these distortions are visible, they seem to have little impact on presence.

Unfortunately, I can offer no concrete explanation for these observations, nor could I find any studies on the perception of space arising from non-central viewing of panoramic or curved images, mono or stereo. Held and Banks (2008) demonstrate that a straightforward geometric interpretation of stereo vision does *not* accurately predict the perception of space in planar stereo images when a screen is viewed from an angle, a result that applies to any non-central viewer in the panorama. This result is confirmed by Pollock et al. (2012), who analyse view distortion for untracked viewers in a CAVE, and find perceived distortion to be "significantly less than the predictions of the ray-intersection model based on stereo viewing geometry."

Beyond this, I can only speculate. Perhaps the partial visibility of the screen surface results in a partial rectification of space, although this begs the question as to whether the robustness of perspective is something that can occur in gradations, or is simply on or off. Or it may be that the surface-compensation mechanisms are not the whole story, and there are pictorial-compensation mechanisms that come to the fore under certain conditions. The various monocular depth cues within the image (relative size, texture gradients etc) may allow for the partial correction of distorted stereo cues. What is sure is that the perception of space is not as straightforward as a simple geometric projection, for experience within the AVIE reveals that, even with a stereo image, some manner of spatial rectification seems to be at play. All of this warrants further investigation.

2.5.6 Ego-motion, Vection and the Perceptual Rest Frame

Above, it was asked whether a certain degree of immersion could be sacrificed in return for an increased number of viewers. It would seem that the stereo panorama is affirmation of such an idea. A 10 metre diameter panorama, such as the AVIE, can provide a somewhat immersive experience for up to 30 viewers at a time, and there is nothing to prohibit the construction of arbitrarily large panorama to accommodate any number of viewers. The sacrifice made, however, includes above all the abandonment of the ambulatory (x, y, z) dimensions of the plenoptic function. In place of correctly exhibiting motion parallax, the virtual world now tends to swim and distort as the viewer moves about within the light field, in the manner described above. There is no doubt the absence of these dimensions of vision has a destructive influence on presence (Slater et al., 1998). Can, perhaps, the loss of these dimensions be somehow mitigated?

The answer lies in simulation of ego-motion, achieved by putting the virtual viewpoint in motion through the virtual scene. This has three very important results. First, motion parallax is reintroduced into vision, and with it comes the increased perception of depth and space (Gibson, 1950; Rogers & Graham, 1979, 1982; Ono, 2008; Nawrot & Stroyan, 2009). However, as mentioned previously, there is a difference in the perception of space between an actively self-propelled observer and a stationary observer, *even when the visual stimuli are identical* (Wexler & Van Boxtel, 2005). The difference arises, at least partly, from the extra-retinal motor and proprioceptive signals triggered by the moving observer. So while a moving viewpoint certainly does increase perception of space and depth, it is not a perfect substitute for real self-induced physical movement.

Second, motion has the curious benefit of masking the distortions introduced by the non-central observer described above. In particular, for a moving observer, the swimming, unstable appearance of the virtual world is dramatically reduced when the virtual world is in motion.

Third, and perhaps most importantly, is that *a moving point of view induces in the viewer an overwhelming physical sensation of being in motion oneself*. Note, this is by no means a trivial result, for it is perfectly normal to watch images produced by a moving camera, as on the television or in a cinema for example, without feeling that oneself is in motion.

It would seem that the AVIE is capable of inducing *vection* – the illusion of self-motion induced purely by visual or aural stimuli (Fischer & Kornmüller, 1930; Tschermak, 1931). This is an important result, for self motion, or ego-motion, is intimately linked with presence.[57] Riecke and colleagues, who study the link between spatial presence and vection in detail, suggest "that a realistic perception of ego-motion in VR is a fundamental constituent for spatial presence and vice versa" (Riecke et al., 2006, p. 1).

To understand the connection between ego-motion, vection and presence, it is useful to introduce the concept of the *perceptual rest-frame* (Prothero, 1998, p. 27). To paraphrase Prothero, we have a strong tendency to perceive certain elements of our perceived environment as stationary. For instance, the ground or the walls of a room are typically perceived as immobile, and we perceive any relative motion between some object and the ground or the walls as implying that it is the former that is in motion, while the ground or walls are at rest. That is, the walls and ground are adopted by the perceptive system as a reference frame by which all other motion is judged, and to be at rest is to be at rest relative to this reference frame. Gibson was evoking a similar idea when he wrote:

> Just as a motion for the physicist can be specified only in relation to a chosen coordinate system, so is a phenomenal motion relative to a phenomenal framework. Perceived motion occurs in a perceptually stable space or environment. Another way of saying this is to assert that the perception of stability is part and parcel of the perception of motion; you cannot have the latter without the former. (Gibson, 1954, p. 310)

[57] Gibson once noted that perception of one's environment was "inseparable from the problem of ego and its locomotion" (Gibson, Reed, & Jones, 1982, p. 394).

When there are multiple reference frames on offer – and any visual entity may serve as a potential reference frame, including one's self - the perceptive system will tend to automatically select one as the dominant rest-frame, and this shall be perceived as being at rest. This selection is not only automatic, but beyond our conscious control. For example, despite knowing differently, it is impossible to not see the Earth as standing still while the Sun inches across the sky.

Vection occurs when some object in motion relative to the observer, (be it virtual or real) is perceived as a rest-frame, giving rise to the illusion that it is the self that is in motion. A familiar example is when we are seated on a stationary train, perhaps awaiting departure at the station, when a neighbouring train begins to pull away. It is not uncommon for this to induce a very strong sensation that the other train is stationary and it is our own train that is moving, in the opposite direction. This is vection. The illusion is often only temporary, and at some point the rest-frame snaps back to our own stationary train, demonstrating the potential volatility of the selected rest-frame when multiple candidates are on offer. A similar sensation can be gained gazing down from a bridge upon a flowing river. A more dramatic example is Amariah Lake's *Haunted Swing* illusion, a 19th century fun-fair attraction, in which the walls, floor and ceiling of a room are put in dramatic motion around a stationary swing structure, giving the very strong impression that the swing is in motion (Lake, 1893). Wood (1895) recounts the very physical and irrepressible nature of the vection induced by the device:

> Each vibration of the swing caused those peculiar "empty" sensations within which one feels in an elevator; and as we rushed backwards toward the top of the room there was a distinct feeling of "leaning forward," if I can so describe it - such as one always experiences in a backward swing, and an involuntary clutching at the seats to keep from being pitched out. ... The curious and interesting feature however, was, that even though the action was fully understood, as it was in my case, it was impossible to quench the sensations of "goneness within" with each apparent rush of the swing. ... the sensation before described was always present (and I visited the place several times), though I tried to suppress it and reason against it. (Wood, 1895).

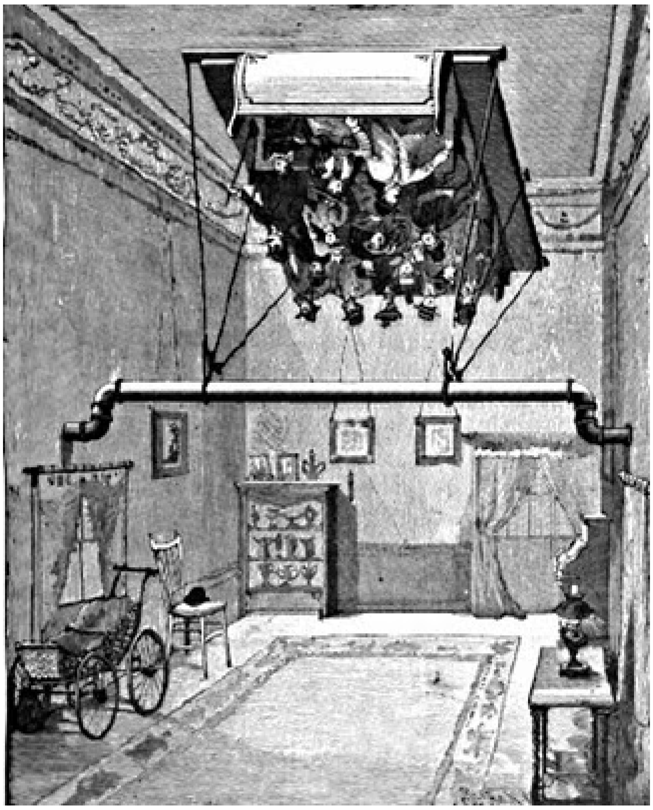


Figure 7 - The Haunted Swing Illusion (1893). "The device was worked in the following way: The swing proper was practically at rest, merely being joggled a trifle, while the room itself was put in motion, the furniture being fastened down to the floor, so that it could be turned completely over." (Wood, 1895). Image from Hopkins (1898, p. 92).

In an immersive display such as the stereo panorama or a CAVE, the viewer is presented with *two potential rest-frames*: the real world or the virtual world. Presence, Prothero suggests, can be equated with the *degree to which the virtual world influences the selected rest frame* (Prothero, 1998, p. 31). In other words, the world that we perceive as being at rest is the world within which we are present. This is a very elegant framing of presence, for it anchors presence to a very concrete and measurable feature of perception, and provides a mechanism for understanding the transfer of presence from one world to another.

> Presence reflects selected rest frame decisions. This suggests that it is possible to measure presence by creating a conflict between real and virtual rest frame cues and then evaluating the relative influence of the virtual cues on the selected rest frame. Thus, a scale for presence can be constructed in terms of the ability of a virtual environment to perceptually overwhelm conflicting real stimuli. (Prothero, 1998, p. 37)

The relationship between presence and vection is most likely more nuanced than this, for there is evidence of a certain bi-directionality between the two. That is, while presence may be affected by, or a reflection of, the strength of the illusion of self-motion (Prothero et al., 1995; Riecke et al., 2004), the perception of self-motion has also been shown to depend on the degree of presence (Riecke, 2006). Further, the concept is only useful for virtual experiences involving motion. Nonetheless, the concept of the perceptual rest-frame presents such a simple and surprisingly powerful device for reasoning about presence that it seems right to explicitly include it as among the necessary features of an immersive display listed above in Section 2.5.1. To reiterate: an immersive display must permit the perception of the virtual world as the *dominant perceptual rest-frame*. Consequently, in order to evoke presence in a virtual environment, it is necessary to understand the factors that influence vection.[58]

The factors that determine the on-set of vection are numerous, complex, and the subject of a great deal of study.[59] (Riecke & Schulte-Pelkum, 2013) divide such factors into two categories: low-level "bottom-up" perceptual factors and higher-level "top-down" cognitive factors. The former includes field of view, velocity, density and contrast of visual stimuli, motion jitter and the curvature of the trajectory. The latter includes such things as perceived foreground-background relationships, ecological validity and realism, plausibility and the conscious knowledge of the possibility of motion, attention and cognitive load, interaction and cross-modal stimuli. All of these factors can be employed in the design of an immersive display and immersive experiences.

One of these factors – foreground-background relationships – helps explain why the presence of other viewers in the panorama tends to strengthen perception of space and motion. Vection tends to be dominated by the motion of the perceived visual

[58] If presence must be reduced to a simple, single-dimensional concept, then this rest-frame construct is a worthy candidate. It is certainly provides explicative and prescriptive powers beyond that of the principle of *celare artem* or the "disappearance of the medium."

[59] The study of vection dates back 150 years (Von Helmholtz, 1866; Mach, 1875; Warren, 1895; Urbantschitsch, 1897). For a comprehensive review see (Dichgans & Brandt, 1978). For an overview of vection in virtual environments see (Hettinger, 2002; Riecke, 2010; Riecke & Schulte-Pelkum, 2013).

background, and it has been shown in a variety of studies that foreground objects or occlusions facilitate vection (see Riecke and Schulte-Pelkum (2013, p. 40) for an overview). Having other members of the audience stand between yourself and the screen not only provides additional depth cues such as relative size, occlusion and parallax, but firmly establishes the virtual scene as a visual background, which strengthens the perception of the virtual world as rest-frame.

Figure 8 - Fellow audience members provide additional spatial cues. (Scenario, iCinema 2010).

It is interesting to contrast this to traditional cinema, where great efforts are made to suppress the presence of one's fellow movie-goers from the conscious mind as far as feasibly possible. The darkness of the auditorium, the tiered rows of fixed forward-facing seats providing, hopefully, an entirely unobscured view of the screen, and a strict moratorium on communication of any kind, all contribute not just to a sensation of being alone, but to a sensation of being *nowhere*. Further, the passiveness of viewing, remaining inert for many hours at a time, adds to the sensation of having *no body*. Traditional cinema, it seems, is predicated on the erasure of one's physical surroundings, one's companions, and one's body from the conscious mind. If the movie-goer does enjoy a form of presence in the virtual world of the movie, then it is a presence predicated on being bodiless, and therefore quite unlike the spatial presence discussed here.

Figure 9 - Fellow audience members provide additional spatial cues. (Scenario, iCinema 2010).

Spatial updating

Could the *illusion* of self-motion provide similar benefits to that of *actual* self-motion? It seems at least partly so, for there is evidence that the illusion of self-motion does indeed improve *spatial updating*. When we move through space we automatically and subconsciously maintain a mental picture of our spatial environment and our position within it, and it is this spatial updating that allows us, for example, to accurately point to objects around us with our eyes closed. In addition to hearing and vision, a wealth of physical cues - proprioceptive, somatosensory, vestibular - play a role in this spatial faculty. In fact, these physical cues play a dominant role, and alone are sufficient for accurate spatial updating, a fact that is well demonstrated by our ability to walk about blindfolded and still point accurately towards objects in the environment (Klatzky et al., 1998). When we are deprived of these physical cues, such as when the viewer remains stationary and motion is represented purely visually, this automatic and obligatory spatial updating of our location within the environment suffers greatly. In other words, our mental picture of our spatial environment is very much tied to our *physical* locomotion through it.

However, it has been demonstrated that if the visual or aural stimuli are sufficient to induce a compelling illusion of self-motion, then spatial updating is significantly improved, approaching the accuracy normally enjoyed with natural physical motion (Riecke et al., 2005; Riecke et al., 2007; Riecke, Feuereissen, et al., 2012; Riecke, Sigurdarson, et al., 2012). This is an important result, for it suggests that at least some of the benefits of physical movement can be reclaimed by fooling the body into believing it is in motion.

Curiously, the potential for vection to improve spatial updating is different for linear and circular vection. (Klatzky et al., 1998) demonstrates that visually induced linear vection – i.e. translations – reliably triggered spatial updating, but circular vection – i.e. rotations – sometimes failed to register mentally. That is, people would maintain an inner mental map of the world as if they had translated in space, but never turned. However, if the viewers physically performed the rotations, but only virtually performed the translations, spatial updating would once again be perfectly accurate (Riecke, Feuereissen, et al., 2012).

The results are still tentative, but if true, the ramifications for the panorama are significant. For when motion is displayed in the stereo panorama, *translations are virtual, but rotations are very real*. As the panorama presents a complete 360° view, and viewers can physically turn to face any direction they like (which is almost always the direction of motion), there is seldom any need to *virtually* rotate the scene. Note, this is not to say that the trajectory of the AVIE through virtual space is restricted to straight lines. The AVIE can follow any manner of curved path, but ideally the orientation of the AVIE should remain fixed, inviting the viewers to physically orient themselves with the trajectory of flight, if they so desire. In this way, the stereo panorama is an ideal arrangement for the inducement of spatial updating through a combination of linear vection and physical rotations.

2.5.7 Interactivity and Presence

It is necessary to expend a few words on the subject of interactivity, for the topic arises frequently in presence literature, and it will play a role in the development *of La Dispersion Du Fils*. A little care must be taken, however, as the relationship between interactivity, immersion and presence presents an inexhaustible field of inquiry, such that even the most cursory consideration of the topic can quickly become overwhelming. As such, I will restrict the discussion to just a few points concerning the role of interactivity in immersion and presence.

When people speak of interaction with a virtual world, they typically have in mind a form of causal interaction, where the immersant's actions causally influence the unfolding narrative of the virtual world. Such interaction includes the manipulation of objects, communication with virtual agents, firing weapons, killing, building and so on. That is, any interactivity that bestows the viewer a form of causal *agency* within the virtual world. It is this causal interactivity that is often identified with increased presence. For example, Kwan Min Lee writes

> If users can make changes to objects that they are perceiving, manipulation, a higher level of experience, occurs. For example, changing the location of an object in a virtual environment is a higher level of experience than the mere act of perceiving the object. When users and experienced objects mutually affect each other, the domain of user experience goes beyond the physical world and an even higher level of experience— interaction—occurs. (Lee, 2004, p. 34).

Such a picture of interactivity and presence could not be further from the position held here. Rather, the "mere" act of perceiving the object is held here as the very foundation of presence, and the ability to effect change in the world is only relevant in so far as it influences our perceptions. This evokes the maxim of perceptual control theory: "behavior is the control of perception" (Powers, 1973). That is, interaction, with respect to presence, should be viewed not as effecting change in the world, but as effecting change in perception.

Interactivity serves presence when it facilitates active-perception of the world, and the most basic form of interaction with the world is the navigation of one's plenoptic function. This raises a small issue of terminology: for whether or not pure ego-motion is considered 'interaction' is sometimes ill-defined. In the real world, when I walk around my house without touching anything, I do not normally describe myself as interacting with my environment. On the other hand, were one to compare a first-person video of a tour of my house with a real-time simulation of the same, in which the viewer has control of their path, one would naturally call the former non-interactive and the latter interactive. That is, with respect virtual environments, ego-motion is generally considered a form of interaction, and it is perhaps useful to speak of different forms of interaction: ego-motion, manipulation of objects, communication, and so on.[60]

The impact that interactive manipulation of objects has on presence can be understood using very much the same concepts introduced above. When we manipulate an object, we actively create and detect invariant structures across multiple modalities (hearing, touch, sight) and from cues arising from within (efference) and without the body (ex- and re- afference). Our actions cause perceptions, which cause actions and so on. Clearly object manipulation is a form of active perception, and therefore a form of presence.

This has two implications. First, it shows that once active-perception is recognised as a form of interaction, then the importance of interaction with respect to immersion cannot be overstated: without interaction, immersion is impossible. However, this is just a less precise way of saying that without active perception, immersion is impossible, and should certainly not be interpreted as meaning that causal interaction is a necessary ingredient for presence.

Second, we begin to see that active-perception, and therefore presence, is not confined solely to ego-motion. This implies that it may be possible to construct immersive experiences based, for example, solely on manipulation of objects. In fact, as will be seen below, it implies that *any* system in which we actively construct and detect invariant structures in sensorimotor stimuli may support presence.

Causal Perception

The focus on interaction as active-perception does not deny causal agency any role in presence. Rather, it highlights an interesting aspect of causal agency: it is only the

---

[60] The term is often used inconsistently. For example, actively perceiving a static hologram would commonly not be considered interaction, but if the exact same vision of the world were achieved using an image display and using a mouse to move viewpoint, it probably would. And if the image display were equipped with head-tracker, would this be considered interactive?

perception of agency that influences presence. It is only when our actions give rise to percepts, and the action/percept bindings are perceived as cause and effect, that causal interactions constitute a form of presence. It is not sufficient to know that one's actions are causing effect in the virtual world; rather they must be *perceived* as such.

The perception of causality, rather than a higher-level cognitive process of induction or inference, has been revealed to be a surprisingly low-level perceptual faculty (Michotte, 1963). The perception of two co-occurring events as causally linked is subconscious, "phenomenologically instantaneous, automatic and largely irresistible" (Scholl & Tremoulet, 2000; Scholl & Nakayama, 2002). For example, when a moving object suddenly stops next to a stationary one, and simultaneously the stationary object springs forth, we irresistibly perceive the former as causing motion in the latter, and perceive the two motions as a single unified percept: a collision. Michotte explored this aspect of perception extensively, demonstrating that the perception of causality is governed by highly constrained law-like relationships between stimuli. This is evocative of Gibson's invariant structures, which too describe law-like patterns in stimuli.

Furthermore, perception of causality seems to be yet another form of multi-modal spatiotemporal *binding* of stimuli, in which stimuli caused by multiple events (the two objects) are bound together to yield the perception of a single event (a collision between objects) (Buehner & Humphreys, 2009; Buehner, 2012). Like other examples of binding, such as the precedence effect or summing localisation (see Section 3.10.1), it is possible, through experimentation, to delineate the necessary conditions for binding to take place. With causality, it seems one key factor is the spatiotemporal proximity of the constituent stimuli. For example, by introducing a delay as little 150ms between two events, the perception of causality can be destroyed, even when the events are known by the viewer to be causally linked (Michotte, 1963). An example of this in the real world is watching someone from a great distance cut wood or hammer nails. Even when the loud striking sounds are known to be caused by the blows of the axe or the hammer, because of the sonic delay, it is impossible to perceive them as such.

In addition, like other forms of binding (for example, the ventriloquism effect), causal perception has the potential to alter other perceptions. Buehner (2012) shows that a perceived causality can result in "a subjective shortening of elapsed time between actions and their resultant consequences." Kim et al. (2013) show that a perceived collision can alter the perceived trajectory of motion *leading* to the collision. They conclude that

> perceived causal relations among visual items are not merely a summary interpretation imposed on motions already determined by perceptual processes, but rather may make a potentially fundamental contribution to the disambiguation of the underlying sensory signal itself. (p. 7).

These examples can all be considered evidence of the low-level nature of causal perception. They also show how causal perception, when considered as the formation of multi-modal invariants through binding of stimuli, is very well accommodated by the model of presence and immersive display adopted here.

The position adopted here is that causal interaction may be considered a form of presence*, but only when it gives way to causal perception*. As will be seen below, when this is not the case, interaction can very easily be detrimental to presence. This position finds support in the work of Cavazza et al. (2007), who study explicitly the connection between presence and causal perception. By measuring presence (by questionnaire) of subjects immersed in virtual worlds in which the laws of cause and effect are artificially manipulated, the viewers reported a correspondence between the strength of their causal perceptions and presence.

Immersive Interfaces

When the viewer looks about within a panorama, they are exercising a sensorimotor faculty directly and without mediation. In contrast, when the viewer looks about with a head-mounted display, their head orientation must be tracked, fed into the system and used to generate the anticipated stimuli. The first is an example of direct interaction, while the second might be described as indirect interaction.

There is a third way, which I shall call re-directed interaction, which involves the use of one physical faculty to control another. For instance, rather than controlling our x, y, z or θ, ϕ coordinates by moving or turning the head or body, these sensorimotor faculties might instead be controlled with hand-manipulated input devices. This is an approach common among first-person computer games, where control of our direction of view (θ, ϕ) is re-mapped to the mouse, and control of (x, y, z) is re-mapped to the keyboard; a re-direction that has become a de-facto standard for a whole class of computer games.

The effect of re-directed interaction on presence is not straightforward. On one hand, the usual law-like relationships between action and perception are missing. When we move the viewpoint to the side with the press of a key, the proprioceptive cues that would normally arise from movement of our head or body are completely absent. Further, the physical cues arising from our stationary head and body are now a source of perceptual conflict. The virtual world spins even when our head is held still, and when we do turn our heads to the side, we are met with incoherent stimuli. Recalling once again Harvey and Sanchez-Vives (2005), for presence to persist, "the constellation of sensorial cues in a virtual environment must be in accord with some basic rules which, in the real world, govern the relationship between sensory events." Viewed like this, it is easy to see how such re-directed interfaces are incompatible with presence.

On the other hand, pressing a key is still a physical action, and therefore still gives rise to a variety of efferent and reafferent cues. Assuming the relationships between these cues are law-like, could not these relationships be learned? Biocca et al. (2001) suggest that when confronted with incoherent stimuli, three outcomes are possible: a) certain stimuli may be suppressed, and/or b) simulation sickness or discomfort may arise, and/or c) an adaption or recalibration of a sensory or motor modality may take place. It is this last possibility that is of great relevance to presence. If, over time, the body adapts to new law-like correlations between reafference and efference, and learns also to suppress conflicting cues, then it is possible that an 'unnatural' interface can be 'learned'.

The learning of new sensorimotor relationships is, in fact, common place. For example, when we learn to ride a bicycle or play a musical instrument, we forge new bonds between action and perception. However, neither of these examples demand the suppression of well-established sensorimotor relationships. That is, our normal sensorimotor relationships are *augmented* by new ones, rather than subverted, and when we turn our head aside, our vision still responds accordingly. Is it possible to 'unlearn' or suppress temporarily basic sensorimotor relationships such as the relationship between movement of one's head and position within the plenoptic function?

Such an idea is dramatically demonstrated in the extraordinary "inverted-vision" experiments first performed by George Stratton in 1897 (Stratton, 1897), and subsequently repeated many times since (Kottenhoff, 1961; Melvill Jones et al., 1988; Gregory, 1997). Subjects wear glasses that invert their vision (in most experiments vertically, but in some, horizontally), and are left to go about their daily business. At first, as might be expected, their sensorimotor faculties are greatly depleted, and the simplest of tasks pose great difficulty, from reading and writing, grasping objects, walking or even reading a clock. However, subjects very rapidly adapt to their new relationship with the world, and within just half an hour, their ability to perform basic actions is dramatically improved. Linden et al. (1999) report that by the third day, the subjects were capable of "walking freely, and performed all tasks of everyday life with none or minimal aid," and by the fourth or fifth day, "they were able to find their way in a crowded department store and to ride a bicycle."

These experiments can be taken as evidence for the plasticity of sensorimotor mappings. They suggest that an unnatural immersive interface, one in which one physical modality is used to control another, can indeed be 'learned', meaning that the relationships between the various efferent and reafferent cues arising from active perception of the world can be learned, and other relationships suppressed. According to this theory, it should be perfectly possible to immerse a viewer with a head-mounted display, but rather than attach the position and orientation tracker to the head, attach it to their hand instead. In theory, after a period of adaptation, presence would ensue.

With this last example, one might suppose that for presence to survive, the control of vision afforded by movement of the hand would need to be as nuanced and delicate as that normally afforded by the head. This brings us to the real challenge facing the design of immersive interfaces: the re-mapping of physical faculties without *loss of accuracy, nuance or expressive power.* Just as the display must preserve the invariant structures and permit the creation and detection of new structures, so too must the interface.[61] When, for example, ego-motion is re-directed through a keyboard, the loss of physical expression is catastrophic for presence. The reduction of physical motion to a handful of binary switches represents such a dramatic simplification of ego-motion that it is difficult to see how sensorimotor bindings sufficiently rich for presence might survive. This is particular so when we consider the importance of continuums and gradients to perception, all of which are extinguished by the on/off nature of the keyboard.[62]

The concept of sensorimotor adaptation suggests that the requirements for presence might be relaxed. While for active-perception to take place, there must exist a law-like relationship between sensorial cues, these nomological relations need not mirror exactly those that arise in the real world. It may be that these relationships need only possess certain key characteristics, such as continuity, determinism and causality. This suggests that we might exhibit a form of presence when we interact with *any* dynamic system that exhibits these features. One such example might be the playing of a musical

[61] Computer interfaces are commonly implemented as a confederation of distinct mono-directional devices, each of which can be neatly classified as either input devices (keyboards, joysticks, cameras, trackers and sensors) or output devices (displays, motion controllers or audio systems).* Although, the word 'interface' is sometimes used in reference to input devices alone, it is far more useful to use it in reference to the complete ensemble of input and output devices. Used like this, the term 'immersive interface' better captures the interconnected roles of action and perception in the evocation of presence. And a *multi-user* immersive interface implies not only input devices that are concordant with multiple simultaneous users, but output devices also. (*There are some rare devices that provide simultaneous input and output, such as a force-feedback actuator, which not only measures force, but exerts force. Or the coupling of an eye-tracker and screen, in which input and output are rendered inseparable.)

[62] This is challenge facing the inducement of presence through object manipulation. To date, there are no interfaces that can provide the subtle combination of haptic, tactile, visual and aural cues that arise from physically manipulating an object. Ego-motion, it seems, is more amenable to re-directed interaction than object manipulation.

instrument, an exercise with a clear law-like relationship between sensorial stimuli, rich in efference, reafference and multi-modal binding. When presence is framed like this, it is easy to draw parallels with Csikszentmihalyi's (1990) concept of *flow*, a state of mental focus experienced by someone performing a complex task with mastery. A musician in a state of flow can be understood as being present in a sonic landscape.

The Perils of Interactivity

An oft overlooked aspect of interaction is that while interaction may enhance presence, *it can just as readily destroy it.* And there are a number of ways it might do so:

- An unfamiliar or unmastered interface not only frustrates and restricts the natural capabilities of the user, but it can force ordinarily subconscious acts into the conscious realm, and give rise to sensorimotor percepts that do not appear to follow known patterns. Further, an unfamiliar interface demands the user focus on the interface (and the real-world) rather than the virtual world.
- An interface might be too 'narrow,' reducing actions ordinarily rich with multi-modal stimuli and motion to acts of extreme simplicity, such as pressing a button. Such simplifications destroy the wealth of multi-modal bindings that normally accompany such acts. Over-discretisation of input (for example, a binary on/off button) has a particular deleterious effect as it destroys any chance for active-perception, in which action and perception must necessarily unfold *over time*. This partly explains the prevalence of computer games in which the user is either firing some kind of weapon, or piloting some kind of vehicle. Both actions are very well suited to the simple 'narrow' interfaces of the keyboard, joystick and mouse. In contrast, games in which the player manipulates objects with their hands are very rare, for the reason that none of the complex multi-modal cues that normally accompany such a basic and familiar action are present.
- An interface not suitably matched to the display can easily exasperate the flaws and limits of the display. For example, an interface that encourages the viewer to walk about physically will instantly expose any lack of parallax in

the display, should parallax be missing. Similarly, an interface not matched to the virtual world can greatly exasperate any flaws and limits of the virtual world.

- ❖ An interface in which the viewer is unable to directly and immediately perceive the effects of his or her actions will be detrimental to presence. This is particularly so in the case of multi-user experiences, where although a user's inputs are indeed playing a causal role in the unfolding events, the user themselves are unable to perceive this causality.

These are just some of the ways interaction may be detrimental to presence. They are drawn from observations and experiences obtained in the AVIE system, and in particular attempts to add interactivity to the work *La Dispersion Du Fils,* as discussed in Section 4.13 below.

2.6 The Art of Immersion

The International Society for Presence Research maintains an active bibliography of presence publications, by no means complete, but nonetheless containing 2770 references as of September 2013 (Lombard & Jones, 2007). The publications are drawn from engineering, psychology, cognitive and computer science, human computer interaction and interface design, education, linguistics, medicine, philosophy, physics, psychology and sociology. Of all these disciplines contributing to presence research, one field is conspicuously absent: the creative arts. Presence research is, it seems, predominately a scientific inquiry.

The absence of the creative arts is striking if one is to consider the proposition that a great body of knowledge about the nature of presence is tacitly incorporated in artistic practice. It is not difficult to find many of the results of contemporary presence research elegantly demonstrated in art, suggesting a tacit understanding of presence, and how it might be manipulated, pervades artistic practice. Mimesis, immersion and presence have been the currency of artists for millennia, and in many ways, artistic practice *is* a form of presence research. The science of immersion was long preceded by an art of immersion, a point central to the writings of Oliver Grau, who argues that "virtual reality is a constant phenomenon in art history that can be traced back to antiquity" (Grau, 1999).

Many of the aspects of immersion and presence presented above can be found elegantly demonstrated in works of art, a handful of which will be discussed here.

2.6.1 The Corporeal Cinema of Cardiff and Miller

One of the more elegant demonstrations of the intimate connection between immersion, ego-motion and perceptual binding can be found in the work *Alter Bahnhof* by Cardiff and Miller, exhibited in 2012 at Documenta 13. The work takes place in the old train station of Kassel. The viewer is given a small hand-held video player and a pair of headphones and directed to sit on a very specific location - a bench by the entrance to the station - before pressing play. What appears on the tiny screen is a film, filmed from a first person point of view and starting at exactly the point the viewer is seated. A voice in the headphones instructs the viewer to physically follow the trajectory of the camera with the little screen as if it were a window rather than a screen and, like this, they are led on a journey throughout the train station.

Figure 10 - Cardiff and Bures Miller, Alter Bahnhof Video Walk, Documenta13, Kassel, 2012.

Despite the simplicity of the contrivance, the effect is extraordinarily immersive. It is a form of augmented reality, but Cardiff and Miller have found an ingenious way to avoid the most difficult technical (and as of today, unresolved) challenges of this medium: tracking the coordinates of the viewer in the light field and blending the virtual seamlessly with the real. With very little effort it is possible to trace *exactly* the path that the camera took with the little handheld screen, to the degree that it appears that one is not looking at a television, but through a camera, and the events are not pre-recorded but taking place before one's eyes.

The effectiveness of this work can be understood in terms of the framework presented above. As the viewer and the camera share the same physical location and environment, there are no conflicts between the real and the recorded perceptual cues that are important to a sense of presence. This is particularly true with the audio, for all the environmental acoustic cues captured in the recorded soundtrack perfectly match the visual environment. All sounds reverberate and reflect, attenuate and diffract exactly as they should, and are duly perceived as happening in the world, rather than in the headphones. At no point is this stronger than at the moment in the film when the viewer finds themselves in a concrete stairwell with the sound of someone rapidly descending from far above in leaps and bounds. With the extremely rich and complex acoustic spatial cues of the locale, the natural visual occlusion provided by the staircase providing a plausible reason for why the descending runner cannot yet be seen, and the viewers own footsteps producing similar sounds, it is, at this point, truly impossible to distinguish the virtual from the real. (Interestingly, listening to the sound of footsteps has been found to activate the same perceptive mechanisms in the brain as those activated when we see biological motion (Bidet-Caulet et al., 2005). This implies that the sounds are not interpreted as the footsteps of someone descending the stairs, but are *directly perceived* as such, in the same manner as Johansson's moving point lights are directly perceived as human motion (Johansson, 1973). The direct perception of biological motion is discussed in greater detail in Section 4.6.5).

The effectiveness of this augmented sonic landscape invites the work to be viewed primarily as a sonic work of art, with the visual component of the film serving primarily as a means of tracing the trajectory of the camera. The film/space is filled with sonic events – dogs barking, wandering musicians, the clatter of rolling suitcases, the arrival and departure of trains – and footsteps resonate at all times throughout the marble-floored Bahnhof. The soundtrack has been recorded using a technique known as *binaural audio*. As discussed in great detail in Section 3.10, much of our ability to perceive the direction and distance of sound depends on subtle changes in intensity, phase and spectra to sound as it travels around and through the listeners head and ears. Binaural audio is simply the capture of these spatial cues by recording with microphones embedded in a physical mock-up of a human head or, in this case, by wearing microphones within the ears. The resulting recording must then be listened to through head-phones (for it is important that the listener's head and ears do not introduce additional spatial cues to those captured in the recording), and when done so, the recorded sounds appear to emanate not from the head-phones, but from their original locations. The technical conceit of binaural audio, combined with the viewer standing in exactly the place the sounds were recorded, eliminates all perceptual conflicts that may inhibit perceptual binding.[63]

The mechanism relies on the hand-held screen only occupying a fraction of the viewer's field of view, for it is necessary to see the real world around the virtual world in order that the trajectory through one is identical to the other. Intriguingly, that the tiny screen occupies only a small fraction of the viewer's field of view seems to have little impact on presence. It is the author's experience that this is unlike other 'moving window' experiences, such as that offered by Jeffrey Shaw's EVE-Dome or PLACE platforms (see Section 2.6.3 below). The reason for this can be attributed to the fact that there is no conflict between the rest-frame within the screen, and that of the real world. And because the viewer has unlimited access to θ and ϕ outside of the little screen by simply looking about, he or she is free to explore the ambient 'panoramic' dimensions of their plenoptic function at will. By contrast, imagine for example that the same experiment were repeated in a completely darkened room. In such a case, it is sure that the extreme limitations on θ and ϕ imposed by the tiny screen would become distinct and obvious impediments to a sense of being there.

As the viewer physically moves the screen about, they are physically engaging in the exploration of their plenoptic function and wave-field. It is a form of physical, corporeal cinema, where the body has been reincorporated into the act of viewing to the

[63] See Simpson et al. (1996) and Gilkey et al. (1999) for further discussion on the impact of concord or conflict between visual and audio cues on presence.

degree that the body is now performing the film. All the ambulatory and ambient dimensions of the plenoptic function are free for exploration.

Now, at this point the reader may suggest that the viewer does not actually have *any* degrees of freedom with which to explore their surroundings. They are bound to follow the trajectory prescribed in the video, and if they deviate from this, the illusion is broken. However, and this is perhaps the most intriguing aspect of the work, it seems that the distinction between cause and effect within the perception/action cycle can quickly become obscured. It seems that the efferent and reafferent cues produced in moving the screen to follow the image are qualitatively similar, with respect to presence, to those that would be produced were the image following a moving screen. After all, the viewer's actions are intentional and their physical movements are no less coherent with their perceptions.

2.6.2 Terra Nova

The immersive theatre works of Belgian theatre company CREW also demonstrate a sophisticated understanding of the mechanics of immersion and presence. For example, their 2011 production, *Terra Nova*, can be considered a study on the impact of multi-modal binding and active perception on presence. Here, the viewers wear head-mounted displays and head-phones, through which they perceive panoramic films and sound, both pre-recorded and live. Meanwhile, CREW members physically manipulate and stimulate the viewer's bodies in tight synchronicity with the events taking place before their eyes. They are tilted and spun, walked around, stroked and rubbed, wet with water and blown with wind. Limbs are physically manipulated and stimulated to provide supporting haptic and somatosensory evidence for the visual and sonic experience.

Figure 11 - Terra Nova by CREW (2011). Photo: Stefan Dewickere.

Figure 12 - Terra Nova by CREW (2011). Photo: Stefan Dewickere.

These theatrical devices employed by CREW can be considered testimony to the importance of multi-modal and conflict-free perception in the invocation of presence.

Figure 13 - Terra Nova by CREW (2011).

2.6.3 Immersion through Interaction

The relationship between interaction and immersion discussed above, in which interaction was shown to be a critical element of immersion in so far as it enabled active perception of the environment, is evidenced in a number of immersive artworks.

For example, in many of Jeffrey Shaw's immersive works, interaction is exclusively limited to the navigation of the plenoptic function, and the works can largely be catalogued according to the degrees of freedom with which the viewer is free to navigate a virtual environment. In the Shaw's *Place* series of works (*Place - A User's Manual* (1995), *Place - Ruhr* (2000), *Place - Urbanity* (2001), *Eavesdrop* (2004), *Place - Hampi* (2006)), an image (sometimes stereo, in other works not) is projected onto the inner-surface of a cylindrical screen from a projector mounted on a central rotating platform. The image acts like a window into a virtual world, and as it pans from left to right, it reveals the corresponding view of the surrounding virtual light-field. The viewer is invited to stand upon the platform and take control of the turning platform, steering it left and right, and thereby taking direct control of their ϕ coordinate in the plenoptic function. A similar arrangement is employed in Shaw's *EVE-Dome* (1993).

Here, a robotic device pans and pitches a projected image across the inner-surface of an inflated spherical dome. The viewer wears a small tracking device on their head so that the projected image-window might automatically follow their gaze across the surface of the dome. Again, by treating this image as a window into a virtual world, the viewer is now free to explore the θ and ϕ dimensions of the plenoptic function. In *Conversations* (2004) the exact same degrees of freedom are given to viewers, although this time they explore their virtual surroundings through the stereoscopic image-windows of a head-mounted display. In *The Golden Calf* (1994) the viewer uses a hand-held display to inspect a virtual object, in this case a false idol, thereby exercising a full 6 degrees of freedom in their exploration of the plenoptic function.

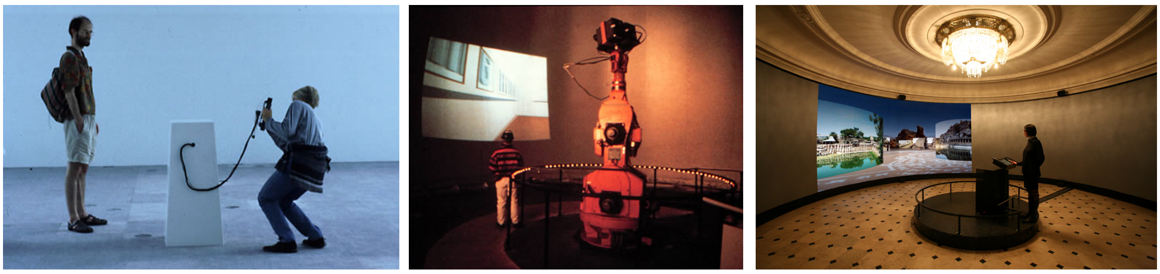


Figure 14 - The plenoptic works of Jeffrey Shaw.
Left: The Golden Calf (2004) Middle: EVE-Dome (1993) Right: Place-Hampi (2006).

All of these works by Shaw allow the viewer control of a virtual window, through which they may peer into a virtual world. He has a made an art of constructing interfaces that permit direct control of these virtual windows, providing natural and intuitive avenues for the active perception of the environment. Further, in all of these works, interaction is very much limited to navigation of space, for the viewer has no more causal powers than controlling their perception of the world. In these works, Shaw employs interaction solely in the service of immersion.

2.6.4 Summary

And so concludes the theoretical part of this thesis. The simple question of the relationship between the panoramic image and sense of "being there," has led us deep into the twin concepts of immersion and presence.

Presence was duly defined as active-perception of one's environment. More specifically, presence lies in the creation, through active-perception of intra-modal and multi-modal invariant structure. This understanding of presence was arrived at by applying J. J. Gibson's approach to ecological perception to the concept of mimetic immersion.

In the following section, this theory of presence will be put into practice, in the construction of an immersive panoramic virtual reality theatre.

3. AVIE: THE ADVANCED VISUALISATION AND INTERACTION ENVIRONMENT

3.1 A CYCLORAMA FOR THE 21ST CENTURY

> The invention that we are about to describe in a general way seems destined, if the hopes of the inventor shall be justified, to bring panoramas into fashion again and to assure them, in the future, new success and a less ephemeral existence. (Chase's Electric Cyclorama, Scientific American, 1896).

This chapter introduces the *Advanced Visualisation and Interaction Environment,* or AVIE, a modern re-imagining of the panorama based not only on contemporary virtual reality techniques and technologies, but on current understanding of immersion, interaction and presence. Taking the form of a cylindrical display, within which the audience view an imaginary vista, the AVIE can be firmly placed within the long tradition of panoramic theatres, from the panoramic cinemas of the 1950s and 1960s - Disney's *Circarama* or the *Krugorama* of Moscow - back to the early extraordinary experiments in panoramic projection at the dawn of the 20th century. However, the AVIE possesses four pivotal features that distinguish it from its ancestors: stereoscopic imagery, spatial audio, real-time image generation and interactivity.

In this chapter, a very brief account of the history of immersive multi-user systems is presented, before the various design constraints and requirements of the AVIE are outlined. This is followed by a detailed technical description of the resulting system and the various strengths and weaknesses of the system, and the extent to which it meets the original requirements and objectives, are then discussed.

3.2 BACKGROUND

In 2002, the iCinema Centre for Interactive Cinema Research, a joint venture between the College of Fine Arts and School of Computer Science at the University of New South Wales, was established with a research agenda primarily concerned with the possibilities and challenges of digitally-mediated immersive and interactive experiences. In particular, the possibilities arising from the application of contemporary immersion and interactive digital technologies (i.e. "virtual reality") to the narrative, visual and cinematic arts. The iCinema Centre's adopted *modus operandi* would be "research by practice," an approach founded on the practical realisation of interactive immersive artworks or, more generally, "experiences," suitable for public exhibition.

In order to pursue this program, a development platform would be needed; an apparatus that would serve as the iCinema Centre's primary laboratory for experiments in immersion and interaction. For this purpose, the *Advanced Visualisation and Interaction Environment* was conceived. As a general-purpose display platform, it would encapsulate all that was common among the envisaged experiments that would follow, with the express purpose of avoiding, as much as possible, the re-invention or re-implementation of tools and technologies with each new project or experiment. Thus conceived, and following the acquisition of UNSW and Australian Research Council grants, in 2004 work began on the design of the *AVIE*.

Amongst the research interests of the iCinema Centre can be found "co-evolutionary narrative" (Kenderdine et al., 2007; Del Favero & Barker, 2010), a formalisation of narrative as emerging from the interactions of *narrative agents*, be they human players or artificially intelligent machine agents. In such a conception of narrative, all interactions between narrative agents are considered equal, whether they take place purely in the virtual world (agent to agent), in the real world (human to human) or across the divide (human to agent). It follows then, that in order to construct co-evolutionary narratives of this sort, an environment capable of accommodating multiple users is required, and that this environment must include the means and mechanisms necessary for interactions to take place, not only between viewers and the virtual world but between viewers themselves.

What was needed, then, was a medium that was simultaneously immersive, interactive and amenable to multiple users. The panorama, a medium that occupies a position midway between single-user immersive systems such as the CAVE or head-mounted display, and traditional multi-user screen-media, such as cinema, would provide the answer. The AVIE would, therefore, take the form of a panorama.

Jeffrey Shaw

It is important to note that this understanding of the panorama was not arrived at *ex nihilo,* but arose directly from the work of Jeffrey Shaw, then director of the iCinema Centre and the primary force behind the creation of the AVIE. The AVIE represents a logical evolutionary step in a legacy of immersive and panoramic platforms developed by Shaw during his long artistic career. This legacy includes the experiments with panoramic cinema, virtual reality and interactive and immersive art conducted at the Zentrum für Kunst und Medientechnologie (ZKM) during the 1990s, where he was director of the ZKM's Institute for Visual Media prior to arriving at iCinema and, travelling further back through his career, the experiments in expanded cinema conducted with Tjebbe van Tijen, Theo Botschuijver and others from 1966 to 1983 (Duguet et al., 1997).

Jeffrey Shaw's work is well known to academia; indeed it would be difficult to find a discussion of media art that does not mention his significant influence on the development of the genre. Rather than repeating what has been said better elsewhere, I leave it to Oliver Grau to summarise his work:

> Shaw is regarded as a pioneer of interactive art. For decades he has been particularly interested in immersion, although he has not stated this explicitly; however, the concept of immersion pervades his oeuvre, from his early ‹inflatables,› his work «Corpocinema» (1967), to his works based on the expanded cinema idea which breaks through the limits of the cinema screen, the various versions of his classic «The Legible City» (1988), a square kilometer of virtual urban space with an architecture of letters as high as buildings that can be crossed by bicycle, his «Extended Virtual Environment,» (1993–1995), and his most recent installations, such as «Place Ruhr» (2000). Visions of future cinematography were assembled in the exhibition Future Cinema at the ZKM, which was co-curated by Shaw. His installation «Place Ruhr» not only links the genres of photography and video with virtual art, but Shaw consciously locates it in the tradition of that dinosaur of media and art history of immersion - the panorama. (Grau, 2004, p. 7).

3.3 PRECURSORS

Philippe Codognet (2003), with the aid of Samuel Edgerton (1991), may have recognised one of the earliest depictions of an immersive display in Giovanni Fontana's 1420 manuscript *Bellicorum instrumentum liber, com figuris et fictitys litoris conscriptus*.[64] Fontana, a Venetian physician and engineer, filled his manuscript with illustrations of a wonderful assortment of devices, from war machines and mechanical demons to musical instruments, fountains, keys and locks and surgical instruments. The illustrations are accompanied by ever so brief annotations in Latin, half of which are encrypted in his private cipher. Fontana, it seems, hoped to both demonstrate his inventiveness and guard his secrets at the same time.

In the illustration on folio 70r, Fontana clearly demonstrates the principal of the projected image, with a depiction of a form of magic lantern.[65] It is, however, the illustration on folio 68v that most draws our attention. Fontana has drawn what appears to be a castle populated with panoramic images. The text, in which one word appears to have been erased, reads "*Castellum umbrarum eo quod in loco obscuro situatur et [?] intra ponuntur et figure umbrate variantes actus suos ostenduntur*." Assuming the missing word to be *lumina*, this might be translated as "The Castle of Shadows, for the fact that it is situated in a dark place, and lights are placed inside and shadowed figures are revealed adopting different motions."[66]

---

[64] "Illustrated and encrypted book of war instruments."

[65] The text reads "Apparentia nocturna ad terrorem videntium" – "A nocturnal apparition for the terror of those who see it", and then, written in cipher: "Habes modum cum lanterna quam propriis oculis vidi[i]sti ex mea manu fabricatam et proprio ingenio" – "You have a method with a lantern, which you saw with your own eyes and which was made by my own hand and ingenuity." The deciphered Latin text is from (Battisti et al., 1984), while the translation from Latin to English is the author's.

[66] Suggestion of "lumina" is from (L. T. White, 1978), while the translation from Latin to English is the author's, with the aid of anonymous online Latin scholars.

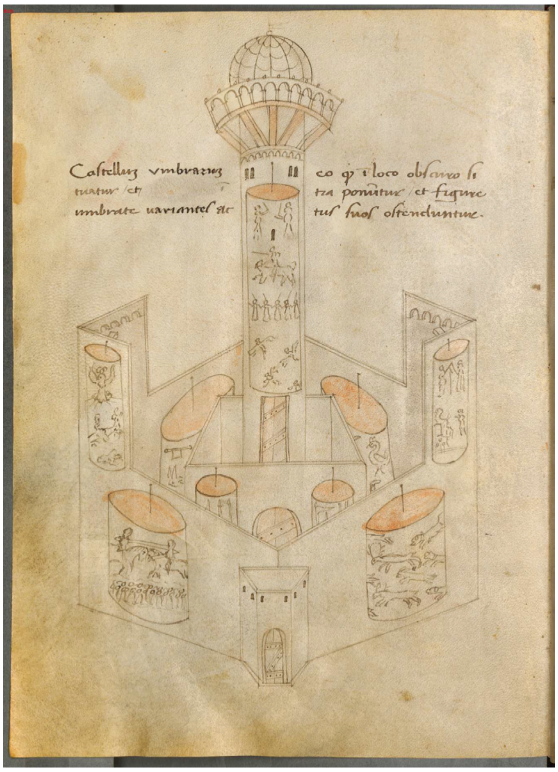


Figure 15 - Castellum umbrarum, Fontana (1420, p. 68v).

Figure 16 - Apparentia nocturna, Fontana (1420, p. 70r).

The cylinders appear to be suspended from above, and thus free to rotate. According to L. T. White (1978), hot air, rising from candles inside, turns light metal turbines at the top of each cylinder, putting them in motion.

Codognet suggests that Fontana’s manuscript offers “a precise description and depiction of a room with walls made of folded translucent parchments lighted from behind, creating therefore an environment of moving images” (Codognet, 2003). This room, he continues, can be seen as an early ancestor of today's *CAVE* systems – an immersive, candlelit cinema of moving screens and flickering images. However, it is difficult to know how literally one should read Codognet’s claims. Fontana’s short cryptic messages fall far short of the “precise description” suggested by Codognet, and while the hot-air turbines described by White do appear in the *Bellicorum*, there is no mention of their use in the *Castle of Shadows*. In fact, closer scrutiny of the illustration of the castle reveals it to be perhaps not an illustration of a real castle, but of a model made from paper. It may well be that the *Castle of Shadows* was not something that Fontana ever intended anyone to enter - for the entire device was no larger than a toy - but to be viewed from without as one does a doll’s house. Nonetheless, it seems churlish to deny that within Fontana's *Castellum Umbrarum* lies the seed of an immersive theatre of projected images.

Electric Panorama

As suggested at the outset of this thesis, the 18th century painted panorama of Robert Barker were very much designed as immersive environments, designed expressly to immerse a multitude of viewers within a single shared vision. As such, these panoramas can be considered direct antecedents of the AVIE.

So too the early experiments in panoramic projection, such as Charles A. Chase's 1895 *Electric Cyclorama*. A device capable of projecting a continuous 360° panoramic image, Chase's design incorporates many features that remain viable today.[67] For example, his system used an array of 8 (or 11) projectors arranged in a circle to project on the inner surface of a cylindrical screen. To achieve a single seamless image, the frusta of the projectors intentionally overlap, and semi-opaque masks placed in front lenses are used to achieve a constant brightness of image across these blend regions, a technique known today as "edge-blending." Chase's *Electric Cyclorama* was equipped with two sets of projectors so that the operator might prepare a new set of slides without disrupting the panorama currently on display. Only once the new panorama was thoroughly prepared would the projectors be masked and unmasked in unison - a technique known in computer graphics as "double-buffering."

In 1898, Thomas Barber exhibited his *Electrorama* in London. Adopting a similar manner to Chase, Barber employed ten projectors to display panoramic still images. However, aside from the enormous dimensions of the screen (12.2 metres high and 122 metres in circumference), little information seems to have survived.[68]

A different and highly inventive approach to the projection of seamless panorama is demonstrated in Auguste and Louis Lumière's 1900 *Photorama*. Twelve lenses are fixed in a ring around a single cylindrical image. However, rather than use overlap and blending to achieve a single seamless image, the entire ring of lenses is set in rapid

[67] See Figure 1 on pg. 18 for an illustration.

[68] Source: *www.alanmachinwork.net/Tourism-s-Educational-Origins-Part-2*

cyclic motion around the cylindrical film. The projected cylindrical image remains static as the gaps between the lenses sweep across the screen. When spun sufficiently quickly, persistence of vision would yield a perfectly seamless cylindrical image.[69]

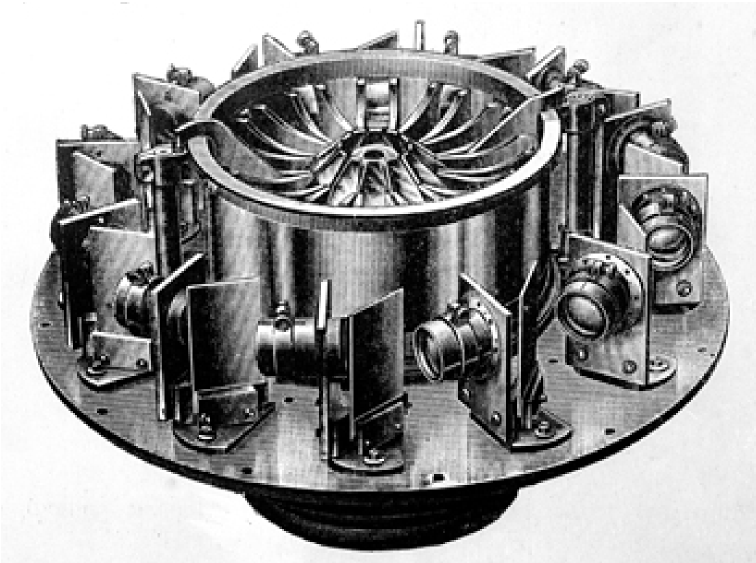


Figure 17 - Lumière's Photorama *projection device (1900), illustrated in Scientific American (1902, p. 344).*

The fascinating aspect of the Lumière's *Photorama* was that, with almost no modification, the system was ideally suited for the presentation of omnistereo panorama. Alongside their projection system, the brothers Lumière had also developed a matching camera, the *Periphote*, which too worked by spinning a lens around a static, cylindrical film surface. By slightly adjusting the angle of the mirror or position of the slit directing the light on to the cylindrical film, an omnistereo image could be captured. Two omnistereo panoramas captured in such a manner, with one lens facing one direction and the other in the opposite direction (either onto the same photographic plate with different colour-filter for each lens, or onto separate plates), would yield matching left and right omnistereo images. Projection of these stereo panoramas could be easily achieved by simply using two rings of projectors, one above the other, or if anaglyphic stereo was being used, no change to the projection device would be required whatsoever. Noting that the Lumière brothers had been awarded patents for "a system of stereoscopic cinema" as early as 1900 and had developed a reliable method for colour

[69] The lenses were spun at a speed of 3 revolutions per second. With 12 lenses, the effective 'frame-rate' was 36Hz.

photography by 1903, it would seem that the Lumières possessed within their knowledge all the necessary ingredients for an anaglyph 3D omnistereo panorama capture and projection system, although there is no record known to the author that a system was ever constructed.

Panoramic Cinema

> The circular image that surrounds us is fugitive.
> It is a projection, and, what is more, a cinematographic one.
>
> *(Scientific American Supplement, 1900, p. 20631)*

The early days of cinema, when the nascent medium was still yet unsure of the form it would take, were marked by a period of great experimentation. Cinema was slowly revealing itself to its inventors as a means of communication, a device for telling stories, for creating art or propaganda, or for documenting the real world. For a small number of inventors, cinema presented an opportunity to return to the project begun by Robert Barker over one hundred years prior. This might be what Bazin famously referred to as "total cinema":

> a total and complete representation of reality; […] the reconstruction of a perfect illusion of the outside world in sound, color, and relief. […] an integral realism, a recreation of the world in its own image, an image unburdened by the freedom of interpretation of the artist or the irreversibility of time.
>
> (Bazin, 1967, pp. 23-24)

The first recorded instance of a panoramic cinema is perhaps Raoul Grimoin-Sanson's *Cinéorama*, in which he employed 10 cinematographic cameras to produce a complete panoramic moving image. In a similar manner to Chase, Sanson's system also achieved a seamless image using edge-blending. The 10 projectors were kept in perfect synchronization mechanically, solving a problem that continues to demand attention today (Grimoin-Sanson, 1896b, 1896a, 1897; Scientific American Supplement, 1900). Sanson famously filmed a flight in a hot-air balloon with a matching 10-lens panoramic camera, marking almost surely the first use of vection as an artistic device in a panoramic theatre. The use of a moving point of view in wide-screen or panoramic images to induce vection would become a standard technique in the panoramic theatres to follow.

To the moving and purportedly colour panoramic images achieved by Sanson, Auguste Baron, in the 1899 patent for his *Cinématorama*, would propose the addition of sound. With a circular array of microphones to capture the sonic environment during filming, and a matching array of loud-speakers during the screening, Baron describes a system for recording and projecting panorama complete with colour, movement and surround sound (Baron, 1897, 1899).

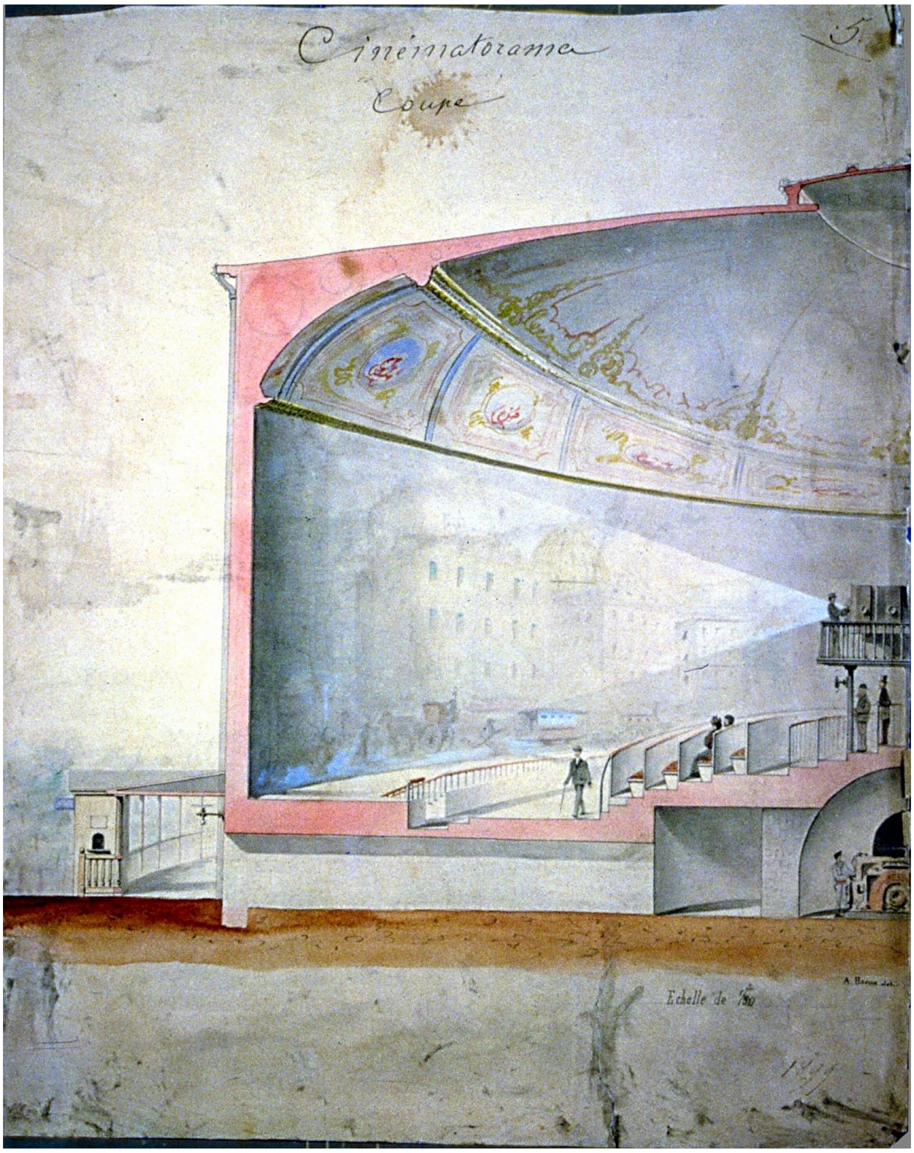


Figure 18 - The Cinématorama of Auguste Baron (1897).

In 1902, Charles Félix André Leguey and Félix Pierre Georges Bap would also file patents for a panoramic cinema camera and matching panoramic projection system

(Leguey & Bap, 1902a, 1902b). Similar to the Lumière's *Photorama*, Leguey and Bap describe a rotating ring of 10 cinema cameras and projectors, in which the movement of film before each lens is tightly synchronized with the rotation of the lenses themselves.

While there is ample record of the panoramic inventions of Chase, Sanson and Lumière in action, it is unsure whether the panoramic devices proposed by Baron or Leguey and Bap were ever successfully constructed. What is sure is that, from the very birth of cinema, attempts to construct the Bazin's "total cinema" abound. Or, conversely, and this is the very point of Bazin's argument, it is from this search for "total cinema" that the conventional cinema of the 20th century was born.

However, even with all the elements for the recording and projection of colour and perhaps even stereo panoramic films being in place as early as 1900, the panorama would not enjoy the "less ephemeral existence" promised by Charles Chase. In fact, following the early experiments at the turn of the century, panoramic cinema all but disappeared, resurfacing briefly in the mid-century, first with Disney's *Circarama* or *Circle-Vision* in 1955, the Moscow *Krugorama* in 1959, and the London *Circlorama* in 1963. None of these systems, however, offered any technical advancement over the techniques of Grimoin-Sanson or Baron. Indeed, with visible gaps between each of the individual projector's images, less accurate temporal synchronicity,[70] and the screen raised above the heads of the viewers, all three systems can be considered a technical regression.

They would, however, confirm the potential of the panoramic image, when captured from a moving viewpoint, to induce vection. All three cinemas would showcase journeys filmed from atop cars, or aboard boats or trains, beneath helicopters, swinging from a crane or sliding across a skating rink. If little else, the films shown in these panoramic theatres represent an impressive catalogue of all the different ways a camera might be put in motion.

---

[70] With the projectors placed outside the cylinder, rather than from a tight ring in the centre, it is no longer possible to keep the projected image in lock-step mechanically and so guarantee a synchronised image, as was the case with the earlier panoramic cinema.

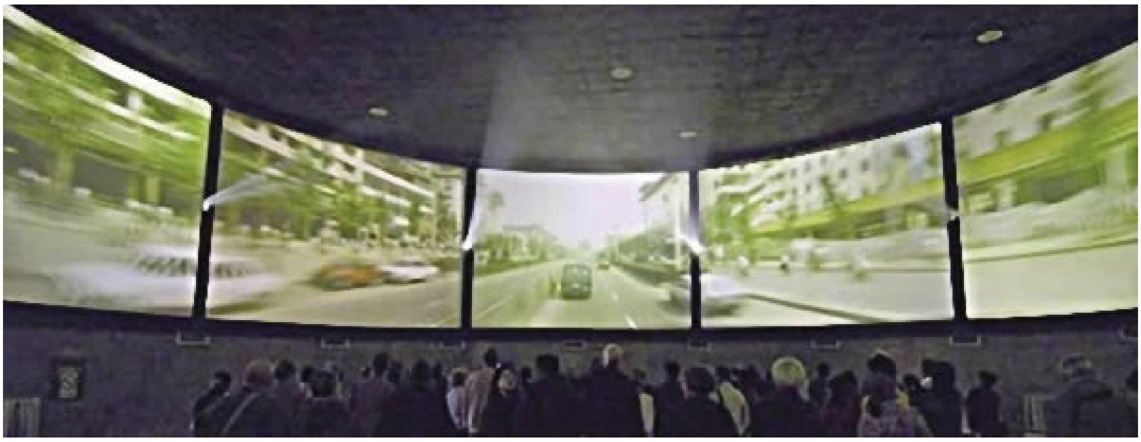


Figure 19 - Disney's Circle-Vision 360 (1955). A 9 camera/projector panoramic cinema.

Figure 20 - Krugorama, Moscow (1960). An 11 camera/projector panoramic cinema.

In 1995, Michael Naimark would demonstrate a novel projection system in which the problem of stereo-panoramic cinema is partially solved (Naimark & Felsenstein, 1997; Naimark, 2005). A conventional stereo-pair of cameras were rotated slowly (1 rpm) to record a 360° scene, and this stereo film would then be displayed on a rotating screen. However, rather than rotate the screen, the floor upon which the audience stands is instead rotated in the opposite direction. The arrangement is successful, for not only does the moving image induce vection, but the viewer will tend to physically walk to keep up with the image, such that the sense of ego-motion is not simulated, but very real. And as the physical act of walking gives rise to a complex array of multi-modal cues, the viewer is physically cajoled into the very kind of multi-modal active-perception that has been identified as presence.

Real-time Multi-User Immersive Projection Systems

In their 1992 landmark paper, Cruz-Neira et al. (1992) report the construction of a new form of theatre. Named the CAVE, the theatre would use an array of projectors to encase a viewer within a cube composed entirely of screens, upon which stereo images are projected. Following Fisher (1981) and Kubitz and Poppelbaum (1973), the location of the viewer's eyes are tracked, and this information is used to render, in real-time, a geometrically correct projection of a surrounding virtual world.

The CAVE successfully demonstrates how images may serve as projective surrogates for the light-field. It can do so, however, for one viewer at a time. To address this issue, Naemura et al. (1998) would introduce the omnistereo projection to the CAVE system, allowing multiple viewers to simultaneously enjoy an imperfect, yet "adequate" perspective on the virtual world.

Recognising that a cylinder was far more suitable projection surface for omnistereo images, Shimamura et al. (2000a, 2000b) proposed the CYLINDRA "stereo panoramic mixed-reality display". With 6 active-stereo projectors illuminating a 330° screen, 6 metres in diameter and 2.4 metres high, a seamless stereo panoramic image with resolution 6144x768 pixels was achieved. Designed in concert with a stereo panoramic camera, the CYLINDRA system could display real-time 3D computer generated imagery embedded in stereo panoramic video streams. Save for the fact that the screen extended only 330°, the CYLINDRA system can be considered the first stereo panoramic projection system.

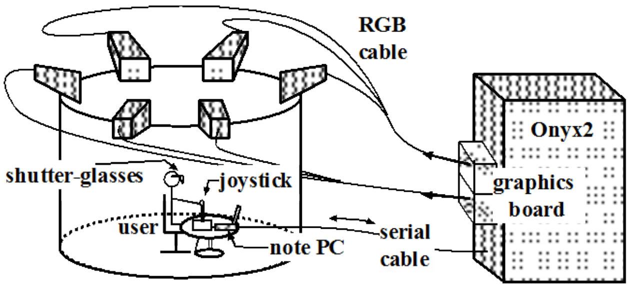


Figure 21 - CYLINDRA, as illustrated in Shimamura et al. (2000a) © 2000 IEEE.

Finally, in 2004, Andreas Simon et al. (2004) used omnistereo to display real-time stereo imagery across a 240° conical screen, successfully demonstrating the use of the omnistereo projection with arbitrary real-time 3D content.

3.4 DESIGN REQUIREMENTS AND CONSTRAINTS

The AVIE would take the form of cinema, or theatre,[71] with a number of specific features. First and foremost, it would possess an interface – a confederacy of immersive visual and sonic displays and interaction devices - conducive to the production of presence. This interface would support up to 20 simultaneous users and would be *non-invasive*, so as to allow, indeed encourage, physical movement and group interaction. It would be relatively low cost and portable. It would be general purpose, a *tabula rasa* like any cinema or theatre, with as little restrictions of the types of experiences or situations it was capable of simulating as possible. It should allow rapid development of new experiments and projects, and it should compare favourably with the state-of-the-art in computer graphics, VR and other interactive media. Finally, it would need to fit within the physical limits of the available space, namely 12m x 12m x 5m. This section presents reasons and ramifications for each of these particular requirements.

Multi-user & Physical

Among the intended applications of the AVIE is the exploration of "co-evolutionary narratives," in which the audience not only interact with the virtual world, but with one another. Hence the requirement of 20 simultaneous users. To aid user-to-user interaction, the interface should be as physically 'non-invasive' as possible. Wires or tethers, heavy or cumbersome input or output devices, or anything that might inhibit

---

[71] The words *cinema* and *theatre* are used throughout this work because, like a cinema or a theatre, the AVIE is a space expressly designed for the mediation of images and sound to an audience. Unfortunately, an undesirable side-effect suffered by both these words is their tendency to conjure images of rows of seats in front of a stage or screen, or worse, the types of narrative experiences that one would normally expect to experience within a cinema or theatre. This is largely a result of the word 'cinema' (and this is equally the case for 'theatre') being a polyseme, meaning at once the space within which one watches a film, the history of cinema, the technology and techniques of cinema, systems of production, economics, markets and its distribution, as well as narratives, styles and cultures of cinema. It is for this reason that the term "immersive cinema" is extremely ineffectual in communicating the potential of the medium. It suggests a direct descendance from modern cinema, inheriting all the customs and traits of its lineage, an inheritance more possessed of constraints than possibilities. Are there better alternatives? "Virtual reality system" is equally burdened with misleading connotations. "Immersive platform" seems free of any historical shadow, but probably summons images of a submarine launch pad before that of an electronic medium.

physical movement, or verbal or gestural communication, are all to be avoided. The interface must not only allow unencumbered movement around the arena, but permit modes of interaction predicated entirely on physical movement or inter-user communication.

Public

A more pragmatic reason for supporting both multiple users and a non-invasive interface is the desire to exhibit the system publicly. The cost of transporting and installing such a large and complex projection system is significant, so it is important to avoid the extraordinarily low spectator-per-hour ratios that typically accompany single-user immersive systems. A solution that allows the public to enter and exit at will without being fitted with special clothing or elaborate equipment, or without instruction or calibration, is far more amenable to public exhibition. These practical considerations impact not just the design of the display, but very much the design and form of any input devices and mode of interaction.

Portability

This desire to publicly exhibit the AVIE also explains the desire for portability. It must be possible to disassemble the system into subcomponents that can be easily transported, while still allowing reassembly at the exhibition site in a reasonable amount of time, i.e. components that are neither too big nor too small. It should be suitably rugged and robust so as to survive repeated installation, dismantling and shipping. Parts should be easily replaced should they fail. Finally, it should not require anything of the installation site other than that which can be reasonably expected of any public exhibition space: sufficient space, clean electricity, and adequate control over the ambient temperature and lighting.

General Purpose

The system should be general purpose, which is to say that it should be free of any innate restrictions on the kinds of experiences that it can mediate.

The pursuit of such generality does however come at a cost. Consider, for example, an industrial flight simulator, where the physical replica of the aircraft's cockpit is key to providing a true-to-life like experience. This high-fidelity is clearly gained at the expense of versatility, for the physical replica of the cockpit precludes the simulation of, say, riding a horse, or a game of golf. In fact, the physical reproduction of the cockpit interior constrains the simulator not only to the experience of piloting an aircraft, but to the experience of flying a very specific model of aircraft.

This example raises a number of interesting issues. First is the conjecture that *all* simulation platforms face an intrinsic compromise between versatility and verisimilitude. If there is a proof to such a conjecture, it might hinge on the notion that, when faced with limited resources (be they computing power, time, space or money), the fidelity of any simulation can be improved by reducing its scope and concentrating more resources on a smaller domain.

Second, although it may seem so at first, generality is not simply a matter of software versus hardware, for both hardware and software are subject to varying amounts of flexibility or constraint. Certainly software has variables, but this is also to say that it has invariables; structures, features and constraints that simply cannot be altered without rewriting the software itself. In the case of the AVIE, this pursuit of generality more significantly impacts the design of the software than the hardware.

Third, in lieu of a perfect immersive interface, free from any display and interaction artefacts that betray the mediation, there will always be some virtual experiences that are better suited to the interface than others. These are the virtual experiences that draw least attention to the flaws in the mediation. A system may be general purpose, but there will be virtual experiences to which it is better suited.

Finally, this example introduces the notion that there is a division between a simulation platform and the simulation itself - a division between 'platform' and 'content.' Where the line is drawn between platform and content is sometimes obvious, but at other time entirely arbitrary. Perhaps the most useful notion is to consider the platform as the sum of the elements of the system that can be easily re-used in different contexts, leaving the rest to be referred to as 'content.' Or those elements of the system that can be easily changed or replaced might be considered the content, while everything that is fixed and invariable can be described as the platform. Either way, the dichotomy of content and platform - the medium and the message - is an intuitively useful concept, even if the boundary between the two is somewhat ill defined.

Space

That the entire system occupy no more than a 12 by 12 metre square and stand no higher than 5 metres tall is a particular constraint arising specifically from the size of the laboratory available at the time of design. Clearly, however, the overall dimensions of the system impacts the logistics of shipping, as well as limiting the choice of locales within which it can be erected.

Content Creation

The production of new interactive immersive content must be relatively easy. This can be interpreted in numerous ways. First, the creation of content for the AVIE should not be significantly more difficult than creating content for an ordinary computer. This implies that the complexity of the AVIE hardware - the multiple computers, the projection system, the sound system, etc - should all be sufficiently abstracted (i.e. hidden) from the developer. This stipulation arises from observations of, and direct experience with, virtual reality systems where the complexity of the system results in a system that very few people are capable of creating content for, save a very rare breed of software engineer equipped with the relevant experience and knowledge. This complexity helps explain why one can observe throughout the VR systems in laboratories and research centres of the world an unfortunately widespread mismatch between the capabilities and potential of the system and the kinds of experiences that are actually demonstrated. There is no shortage of white elephants in the world of high-end virtual reality systems, and it can be difficult for the lay person to understand why a million-dollar facility presents experiences that in many respects seem significantly inferior to those they frequently enjoy on their personal computer or *Sony Playstation*.

One remedy is to ensure the system is compatible with the most widespread techniques and methodologies in the fields of graphics, animation, audio and simulation and so on. If a project requires real-time 3D graphics, for example, then the tools and techniques commonly employed in the computer game industry must be compatible and available for use. This not only helps to achieve standards that compare favourably with the state-of-the-art (which, in this example, is currently set by the computer game industry), but broadens and deepens the pool of artists and programmers who are capable of creating such a project.

Cost

That the system be "low cost" has a number of implications. First, it implies that the theatre be constructed from existing, commercially available hardware, as the development of new computers, projectors or electronic devices is beyond the scope of the project. The use of 'off the shelf' components has become the prevailing trend in VR systems, largely due to the mass production of hardware and software destined for the computer game market. The essence of the concept lies in the idea that, because of the non-linear price to performance ratio of the technologies involved, it is often more financially effective to use a greater number of low-end devices than a smaller number of high-end devices, and still achieve the same overall result. For example, the cost of 4 lower resolution projectors is often significantly lower than the cost of a single projector with the same resolution of the 4 projectors combined. However, the reduction in price comes at the cost of greater complexity, and this has its own financial repercussions. Most notably, greater hardware complexity quickly leads to significant increases in the cost and time of software development.

The total cost of the system also encompasses the cost of maintenance, which includes the replacement of components as they degrade or fail, such as projector bulbs or computer parts or the resupply of 3D glasses to compensate for breakage or loss.[72] In general, however, the cost of replacing malfunctioning components is superseded by

[72] Perpetual costs such as maintenance and upgrades also beg the question of longevity: what might be the lifespan of such a system? Ten years? The first AVIE constructed has been in operation now for around 7 years, and undergone many significant upgrades.

the intermittent need to upgrade components, in order that the facility remains in keeping with the "state-of-the-art." The need to keep in step with the state-of-the-art is linked to one of the culturally determined aspects of immersion. There seems to be an aspect to immersion that hinges on a certain complicity or willingness of the user to look past the obvious flaws in the system, and this selective blindness is largely determined by our previous experiences. For example, it is only after upgrading to a higher resolution display that the pixels in the old display become suddenly obvious and distracting. It seems that as mimetic technologies evolve and such display flaws and artefacts diminish, our tolerance for such flaws diminishes with them, and what was once convincing, becomes unconvincing.

A more significant ongoing expense however, lies not in material acquisitions for repair or upgrade but in the training and retaining of technicians and engineers responsible for the maintenance and day to day operation of the system. And as the complexity of the system increases, or as the technology involved becomes more bespoke, the knowledge and training required becomes yet more specialised, further increasing the expense of retaining the requisite staff. Such costs are particularly acute when the system embarks on a touring exhibition. Here, the cost of freight, while certainly not to be ignored, is rapidly eclipsed by the level of presence required not just during the erection and disassembly of the system but throughout the exhibition.

Software, be it the licensing of commercial tools, or the development of bespoke software, is another ongoing expense, and more difficult to gauge than the acquisition of hardware. The most important budgetary concern, however, is the most unpredictable and often the most overlooked. This is the cost of developing content for the system, without which the system is an empty vessel.

3.5 Geometric Design

Having set out to construct a stereoscopic panorama, the first challenge encountered is the display of a single, seamless stereoscopic image across a large cylindrical surface. Presented here are the various geometric factors and constraints arising when an array of projectors is adopted as an image delivery system.[73]

The AVIE takes the form of a cylindrical screen, within which the viewers sit or stand. The cylindrical form was chosen for its suitability to omnistereoscopic projection, a form of stereoscopic projection that can provide a panoramic stereo image to multiple users. An array of projectors, suspended above in a circular arrangement, projects a seamless panoramic stereoscopic image upon the inner surface of the cylinder. An *array* of projectors is necessary because no single projector (coupled for example with a fisheye or conical mirror lens) would illuminate the entire cylinder with adequate resolution or brightness.[74]

Alternatives to projection do exist, such as a mosaic of LCD digital displays (DeFanti et al., 2009; Navrátil et al., 2009; Papadopoulos et al., 2013). However, such a system would certainly fail the requirements of portability, cost and ease of use. For example, Papadopoulos et al. (2013) uses 416 27" LCD panels to construct the walls of their *RealityDeck*, a four sided rectangular room 10m high, 5.8m wide and 2.8m high. A 10 metre diameter and 4 metre high panorama constructed from the same building blocks would require 550 LCD panels. While in 2013 Papadopoulus and colleagues are able to drive 24 displays with just one computer, at the time of the AVIE's construction 4 displays per computer would have been the maximum, thereby demanding over one hundred computers to drive the system. Clearly, such an approach is at odds with the design requirements set out above.

---

[73] An attempt to be somewhat thorough is made, for while multiple projector systems are commonplace, no precedence for a complete 360° cylindrical stereoscopic system is known to the author. And while the goal is not to reiterate information that can be found elsewhere, it is necessary to devote a certain amount of energy to the explication of some terms, concepts and mechanisms behind projected stereoscopic images, in order that a coherent and logical argument for the numerous design decisions may be presented.

[74] At the time of the initial AVIE prototypes this was certainly true. Today, a single 4096 x 2048 resolution 33000-lumen projector could conceivably be used to illuminate a cylinder with reasonable results (eg. 512 pixels high, 6432 pixels wide at the bottom and 3216 pixels at the top, with 29% pixel efficiency, and therefore 76 lumens/m^2), however the size, weight and sonic noise of such high lumen 4K projectors rules out their use in such a confined space.

Instead, a solution founded on the use of multiple projectors is sought. The use of multiple projectors is a well-established technique for creating display systems with resolutions and fields of view higher than might be afforded by any single projector. The technique is at least as old as Chase's 1894 *Cyclorama*,[75] in which he employed eleven stereopticons (still-image projectors) to furnish the audience with a panoramic view, and was also the basis of Grimoin-Sanson's 1897 *Cinéorama*,[76] where ten synchronised cinema projectors collectively projected a panoramic moving image. In the digital age, the technique grew to prominence with the CAVE (Cruz-Neira et al., 1992), and with the maturation of cluster-based rendering systems it has become the prevailing approach to virtual reality and immersive system design. As such, a large body of technical literature concerning multi-projector displays exists, and the reader is directed to L. P. Soares et al. (2010) for an up-to-date and gentle introduction to the variety of issues and technologies involved.

3.5.1 Front projection

From the outset, the dimensions of the room that would house the first AVIE prototype ruled out the use of *back-projection;* the placing of projectors outside the cylinder so as to project onto the outer surface of translucent screen. Mirrors are a common technique for reducing (folding) the space required for back-projection but even allowing this possibility, no back-projected arrangement was found to be practical given the dimensions of the room. In addition, manufacturing a cylindrical back-projection screen without a single visible seam on either the inner or outer face, and then devising a means to hold it in place without obscuring the image, pose non-trivial challenges that were considered incommensurate with the benefits they offered.

Therefore, a front-projected arrangement was chosen. The projectors would be arranged to project a seamless panoramic image on to the inner surface of an opaque screen. Exactly how the projectors should be arranged is influenced by a surprising

[75] See Chase (1895a, 1895b, 1895c) and (Western Electrician, 1895), (Scientific American, 1896).

[76] See Grimoin-Sanson (1896b, 1896a, 1897).

number of factors, some obvious, others more subtle. In arranging the projectors – in determining their extrinsic parameters (number, position, orientation) – it is necessary to take into consideration the intrinsic properties of the projector (throw and aspect-ratio, brightness, contrast, black-levels, lens shift, depth-of-focus and resolution), as well as the angles the light strikes the screen, the size of the shadow-free zone, the size and shape of the blending regions, the de- or re-polarisation of light upon reflection and the resultant levels and uniformity of resolution, brightness and contrast across the whole screen, and not least, cost. All of these factors enjoy a society of interdependence, so that what follows is a delineation of a process of constraint satisfaction and optimisation, or put another way, compromise.

3.5.2 Projection Distortion

When a digital video projector performs single-point rectilinear planar projection, all light rays pass through a single *centre-of-projection*, which is both the focal point of the lens and the apex of the frustum.[77] The projection is *rectilinear*, in that straight lines are projected as straight lines. In practice, however, it is possible for projectors to possess a small amount of *lens distortion;* an optical aberration producing a deviation from rectilinear to curvilinear projection at the edges of the image, an aberration known commonly as "pillow" or "barrel" distortion. Lens distortion is most prevalent in very wide-angle lenses, or complex zoom lenses, where levels range from 0.5% up to 1.3%.[78] While in most situations such small levels of distortion are imperceptible and can be readily ignored, they do assume significance when attempting to align or overlap images with pixel-level accuracy.

The projection is *planar,* which means that only when the projection surface is a plane perpendicular to the optical axis, will the projected image appear undistorted.

77 The reader is referred to Appendix E for a brief introduction to projector terminology.

78 Distortion levels sampled from the range of projector lenses available at *www.projectiondesign.com*, October 2013.

Conversely, when projecting onto any surface that varies in depth,[79] which is any surface that is not a plane perpendicular to the optical axis, the projected image will be distorted. This implies that no matter how we position our projectors relative to the cylinder, some amount of distortion must occur.

The form of this distortion, which I shall henceforth call *projection distortion*, can be discerned by examining the intersection of the projector's frustum with the screen surface. For example, projecting a rectangle on to an oblique[80] plane distorts the rectangle into a trapezoid or trapezium, which can be seen in the intersection of the four planes that make up the sides of the frustum and the screen surface. This particular form of distortion, which is commonly referred to as *keystone* distortion, is rectilinear, as straight lines remain straight lines. This can be seen by noting that the projection of an arbitrary line on to the screen is equivalent to the intersection of two planes, the first formed by the screen surface, and the second defined by the centre-of-projection and the line being projected. Such an intersection always results in a straight line, thereby confirming that when projecting an image onto a planar surface, regardless of the angle of projection, straight lines will indeed remain straight lines.

What happens when we project a straight line onto a cylinder? Again we observe the intersection of a plane and the screen surface, but this time the answer depends on whether we describe the projected line as a two-dimensional image or as a three-dimensional form. On one hand, the intersection of any plane and cylinder is an ellipse,[81] so it is valid to say that a line projected on to a cylinder is distorted into an elliptical arc. On the other hand, unrolling the cylindrical screen to form a plane rectangle, the elliptical arc becomes a sinusoidal curve (Salomon, 2006; Apostol &

[79] The perpendicular distance measured along the optical axis. The size of each individual projected pixel is not a function of Euclidean distance to the centre of projection (as in the case of a *spherical* projection), but rather is a function of the *perpendicular* distance to the projector, measured along one axis - the optical axis - only. In computer graphics, this axis is almost always labeled *z*, and, when speaking of images, this distance is known commonly as *z-depth* or simply, *depth*.

[80] i.e. any plane whose normal is not parallel to the optical axis.

[81] Result from study of conic sections. e.g. Menaechmus (380–320 BC), Euclid (fl. 300 BC), Archimedes (c. 287 BC – c. 212 BC), Apollonius of Perga (ca. 262 BC – ca. 190 BC).

Mnatsakanian, 2007 for a proof), and so one can equally say that a line projected onto a cylinder produces a sinusoidal curve.

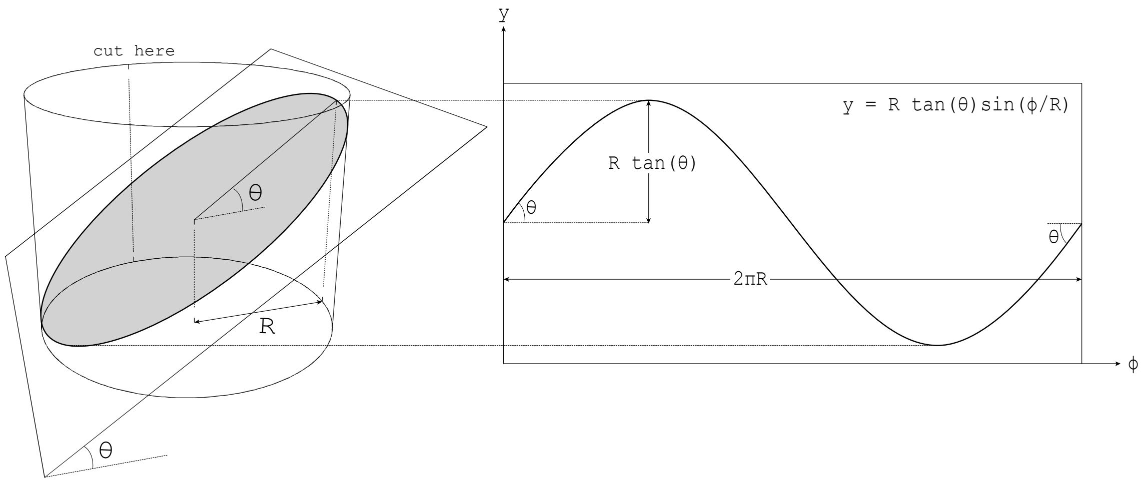


Figure 22 - Intersection of a plane and cylinder. On the left we see the result is an ellipse. On the right, where the cylinder has been unrolled into a plane, the ellipse unrolls to form a sine-wave. The amplitude of the sinewave is $R\tan(\theta)$*, where* R *is the radius of the cylinder, and* θ *the angle with which the plane strikes the screen.*

Both descriptions are correct: described in 3D space, the projected line is an ellipse, while in the 2D reference frame of the screen, or *screen-space*, it is a sine wave. Henceforth, when describing image distortion on non-planar screens - screens that have three dimensions - it is necessary to nominate a frame of reference. This is particularly important when we consider what it means for a projected image to appear undistorted. For example, when a straight line is projected onto a cylindrical screen, what does it mean to say that the projection is undistorted?

From this we learn the exact nature of the projection distortion in the AVIE. To wit, the distortion is curvilinear; straight lines do not remain straight lines, but are transformed into sinusoids/ellipsoids. The four planes that define the sides of the frustum trace sine waves across the screen's surface. Note that the intersection of a plane and a cylinder yields an ellipse/sine wave in all but two cases: either when the plane is exactly vertical (which results in a vertical line), or exactly horizontal (which results in a circle/horizontal line). The planes defining the top and bottom of our projection frustum cannot both be horizontal, therefore at least one of them must trace an elliptical curve/sinusoidal wave as it intersects the cylinder. In fact, as the projectors must be raised above the top of the cylinder, so as to not cast a shadow in one another's light,[82] both the top and bottom planes of the frustum strike the cylinder at an angle. In addition, a slight download tilt of the projector has the sides of the frustum strike the cylinder at an angle, imbuing the image with a small amount of keystone distortion.

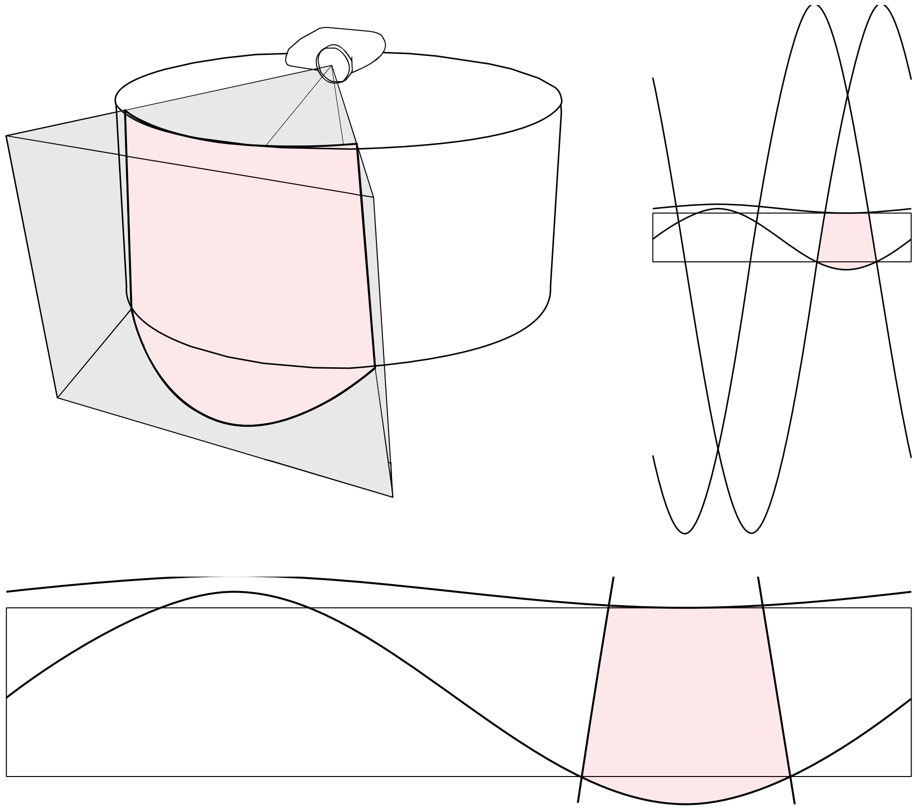


Figure 23 - Intersection of Projector Frustum and Cylinder.
The top, bottom, left and right planes defining the frustum intersect the cylinder to give four elliptic curves, which trace sine waves in screen-space. Together they define the boundary of the projected image, shown here in pink.

While distortion can be corrected, there are a number of reasons for minimising distortion during projection. First, distortion results in non-uniform pixel sizes, which cannot be corrected in software. While non-uniform pixel sizes can be desirable in some special cases,[83] here a globally uniform pixel distribution is sought.[84] Second, following

82 This is unavoidable when the distance from the projectors to the screen is greater than the radius of the cylinder.
83 For example, a head-mounted display might use a non-uniform distribution of pixels in order to exploit the variation of visual acuity across the human retina. The eye has greater resolving power in the centre of vision than it does in the periphery, so a greater pixel density in the centre of vision is highly desirable.

our observation that the top and/or bottom of the projected images must be necessarily curved, we see that it is impossible to completely illuminate the screen surface without having some part of the image spill over the frame of the screen. The greater the distortion, the more light, and pixels, are wasted. *Pixel efficiency* is introduced as the percentage of pixels projected that fall on the screen, and thus contribute to the total resolution and the total brightness.

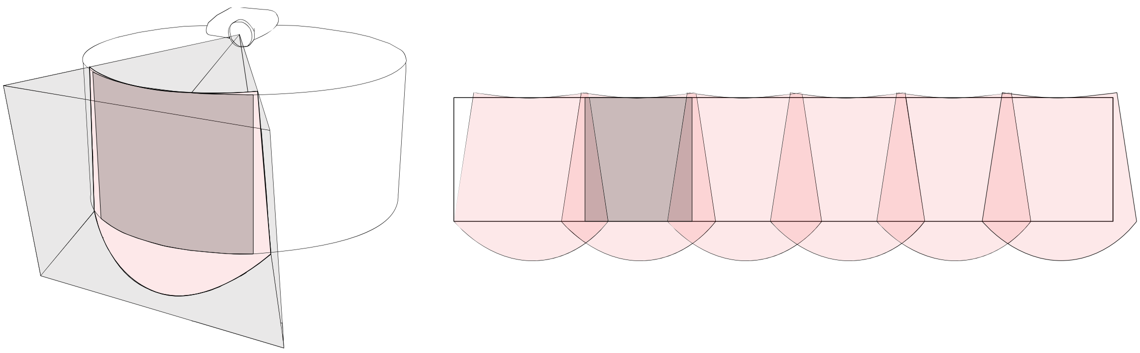


Figure 24 - Aprons on a washing line. Pictured here are the intersections of 6 frustums with the screen.The amount of distortion is exaggerated. All light falling outside the cylindrical screen is lost. Light in the regions where two projectors overlap contribute, on average, 50% to brightness and resolution. The relative size of the grey rectangle to the pink region represents the pixel efficiency – the percentage of projected light contributing towards the image.

Here, pixel efficiency is inversely proportional to the amount of distortion, which in turn is proportional to the amount, *from the point of view of the projector*, the screen surface deviates from a plane. For the cylindrical screen under scrutiny here, this is a function of the radius of the cylinder and the projector's throw, distance to the screen and orientation.

3.5.3 Projector Resolution and Brightness

Projector resolution is one of the few factors considered here that does not require compromise. Put quite simply, the higher the resolution, the better the result. The upper limit beyond which any extra resolution is redundant is given by the maximum acuity of the eye at the closest possible focal distance. Assuming 0.5 arcminute resolution and minimum 10 cm focal distance, this gives us an upper limit of roughly 2,160,000 x

[84] Note that pixel uniformity, like distortion, also requires a reference frame. We can equally describe the distribution and size of pixels on the screen ('in screen-space"), in 3D-space, or, most importantly, across the retina. In this case, uniformity in screen-space is sought.

275,020 pixels for the entire AVIE screen. However, with reduced acuity of the typical viewer, and both shadows and vergence-accommodation conflict precluding viewing the screen from such a close distance, a more realistic maximum resolution would be a fraction of this, but still far beyond current projector resolutions. Any estimate of a minimum resolution depends entirely on the types of images being shown.

Brightness of image too can be unbounded, but attention should be given to two details. First, current DLP or LCD projectors are incapable of projecting perfectly black images, but always 'leak' a certain amount of light, and can only approximate true black with a grey. The brightness of this grey is proportional with the brightness of the lamp, so the brighter the projector, the brighter the grey used to represent black. When projecting a dark image, such as the night's sky for example, these raised black levels are most evident in the regions where projectors overlap, where the 'leaked' light is effectively doubled.

Second, an unavoidable consequence of a cylindrical screen is the reflection of light from one part of the screen to another, and these secondary reflections result in a diminished contrast. As such, there are potential situations (for example, an image of the moon set against the black of night), where a global reduction in brightness can improve overall image contrast and quality.

3.5.4 Projector Depth-of-focus

Another important property of a projector is its depth-of-focus, or depth-of-field. This is a measure of the range of distances over which an image will appear in focus. Conventional projectors are designed to focus evenly on a plane. Therefore, when projecting onto any shape with varying depth, such as our cylinder, variation of focus will result. In a projector, depth-of-focus is inversely related to size of the aperture, and because a larger aperture increases brightness and most projection surfaces are planar, projector manufacturers tend to sacrifice depth-of-focus for brightness. Consequently, depth-of-focus must be considered when determining the optics and arrangement of the projectors.

Like distortion, the range of focus required is entirely a function of the variation in depth of the screen-surface, from the point of view of the projector.

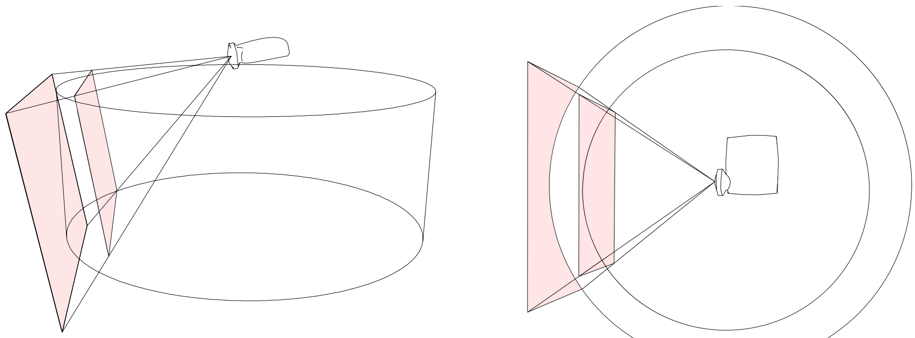


Figure 25 - The two red planes demarcate the range or "depth" of focus required for a sharp image when projecting on a cylinder.

3.5.5 Scattering and Gain

When considering the geometry of the screen and arrangement of the projectors, two angles of significance emerge. One is the angle with which the projected light strikes the screen, the *angle of incidence*, and is a function of the position of the projectors relative to the screen and the shape and dimensions of the screen itself. The second is the *angle of view*, the angle light undergoes reflection before entering the viewer's eye, which is dependent on the viewer's position relative to the screen and, as before, the form and size of the screen. These two angles are important because, for many materials, the manner in which light is scattered, reflected, polarised, transmitted or absorbed when striking a particular surface varies as a function of these two angles.[85] As a consequence, how these two angles change throughout the viewing arena of the AVIE and across the surface of the screen, influences the reception of images in a number of ways.

[85] In optics and computer graphics the relationship between these two angles and the ratio of incoming irradiance and reflected radiance, a relationship which essentially describes how a material appears under different lighting situations, is known commonly as the *bidirectional reflectance distribution function* or BRDF. See (Nicodemus, 1965)

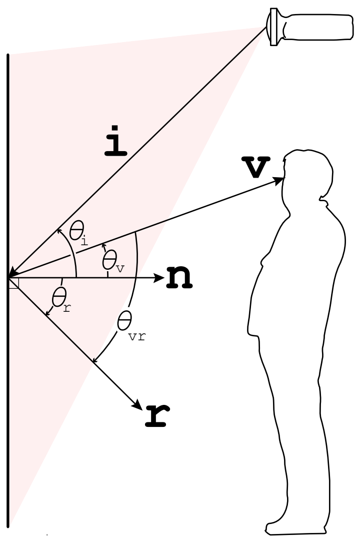


Figure 26 - Angles of incidence, reflection and view.
Four vectors: **i** *- incoming light,* **n** *- surface normal,* **r** *- reflection of* **i** *around normal* **n***, and* **v** *- direction to the viewer's eye. Angle of incidence* θ_i *is the angle between the incoming light* **i** *and the surface normal* **n**. θ_r *is the angle of reflection, and is simply equal to* $-\theta_i$. *The view angle,* θ_v *is the angle of the light striking the eye.* $\theta_{vr} = \cos^{-1}(\mathbf{v} \cdot \mathbf{r})$ *is the angle between* **v** *and* **r.**

First, to understand the effect of varying θ_i alone, it is useful to consider the case of a pure Lambertian surface. A Lambertian surface is any material which scatters light according to Lambert's cosine law (Lambert, 1760), resulting in an apparent brightness that is equal for all view directions. That is, perceived brightness of a Lambertian surface varies with the angle the light strikes the surface θ_i, but not with the angle from which it is viewed θ_v. This is a desirable feature in projection screens, for it ensures an equally bright image for all viewers, regardless of where they are seated.

The brightness of a Lambertian surface is proportional to $\cos(\theta_i)$.[86] From this we deduce two desirable properties. First, in order to make most efficient use of light, we should aim to maximise $\cos(\theta_i)$ across the entire screen surface, which implies the minimisation of angle of incidence θ_i. Second, and perhaps more importantly, in order to attain an image of uniform brightness, we should aim to minimise variation of $\cos(\theta_i)$ across the projection surface. Both these conditions are typically met in traditional cinemas by simply curving the screen around the point of projection, reducing θ_i everywhere to zero. In the AVIE, the cylindrical curve does indeed reduce the

[86] In 3D, it is proportional to the dot product $(\mathbf{i} \cdot \mathbf{n})$.

horizontal component of the angles, but as the screen is without vertical curvature, it is the vertical angles that exhibit greatest variation and extremes.

However, more often than not screen-materials are not pure Lambertian reflectors and do not scatter light uniformly in all directions, but rather exhibit a more mirror-like behaviour, reflecting a greater proportion of light in the direction of reflection **r**, at the expense of other directions. Perceived brightness, therefore, is now a function of one's point of view, with a maximum at $\theta_v = \theta_r$, and diminishing as θ_v deviates from θ_r.[87]

In a traditional cinema setting, where the position of the audience is well constrained, non-Lambertian 'high gain' screen materials - the origin of the term 'the silver screen' for their metallic appearance - are sometimes deliberately employed to make more efficient use of light. By scattering the light less evenly and reflecting a greater proportion in direction **r** towards the audience, a brighter image is attained at the expense of uniformity of brightness. Screen materials are often characterised by two simple factors: *peak gain* and *half-gain angle*, peak-gain being the proportional increase in reflectance at θ_r over a pure Lambertian reflector, and the half-gain angle being the value of θ_v at which brightness falls below 50% the maximum. A 'high gain' screen, for example, might exhibit a peak-gain of 2.4 and a half angle of just 17°.[88]

With non-Lambertian screen surfaces - that is, any screen with a gain not equal to 1.0 - the critical angle is θ_{vr}. It is desirable not only to minimise θ_{vr} in order to make the most of the light produced by the projectors, but to minimise variance of θ_{vr} over the whole screen surface and achieve a more uniform brightness of image.

3.5.6 De-polarization and Ghosting

It should be clear that given the wide range of angles at play in the AVIE, a screen-fabric with a uniform Lambertian reflectance is desirable. Unfortunately this is not possible when using polarised light to achieve stereo separation, as screens that preserve the polarisation of light upon reflection tend to be highly specular or high

[87] Exactly how it diminishes depends on the BRDF of the material.

[88] Data provided by *Pixelution* for their 3D silver projection screen (www.pixelution.co.uk).

gain. High gain screens suffer from poor uniformity of image, especially under wide viewing angles, and lead to "hot spotting;" bright bands or spots when the view angle approaches the angle of reflection. As such, a compromise between preservation of polarity and the uniformity of brightness ensues.[89]

The polarity of light provides a second reason to pay careful consideration to the angles with which light strikes the screen. Whenever light undergoes reflection, it undergoes a change in polarity. This presents a problem when using polarised light to achieve stereo separation, as it is necessary that the left and right eye images remain perfectly orthogonally polarised to one another. Changes in state of polarity result in stereo "cross-talk"; the left and right images are no longer perfectly filtered by the viewer's glasses and each eye now perceives a confusing mix of the left and right images. Exactly how the polarity of light is altered on reflection depends greatly on the physical properties of the reflecting material (rough, smooth, metallic, non-metallic),[90] and the polarity of the incoming light (circular or linear). In this particular context the key observation is that the greater the angles or view, the greater the change of polarity, resulting in a significant increase in stereo cross-talk at high angles (Hong et al., 2010; X. Zhang et al., 2012; Sharp et al., 2013).

3.5.7 Undesirable Curves

Further, all of the above mentioned phenomena caused by high angles of incidence or view are ever so slightly aggravated by the fact that the screen is not a perfect cylinder. When stretching a tube of elastic material between two circular frames, the sides of the cylinder will inevitably bulge inward, forming a shape that can be described

[89] Park et al. (2005) suggest using multiple over-lapping projectors so that every point on the screen is illuminated by at least two projectors. By adaptively changing the source of illumination, on a pixel-by-pixel basis, one can choose the angle of view and avoid specular "hot-spotting." But as this depends on knowing at all times the angle of view, this approach cannot be used with multiple viewers free to move about the theatre.

[90] The changes in state of polarisation exhibited by light undergoing reflection are surprisingly complex. For a detailed account see Können (1985, pp. 144-151).

as a hyperboloid or catenoid.[91] The deformation is not great, not much more than a reduction of 10cm in radius at the midpoint of the cylinder, but nonetheless, as can be seen in Figure 27 below, this curving of the cylinder only serves to increase the range of angle's incidence, view and θ_{vr}.

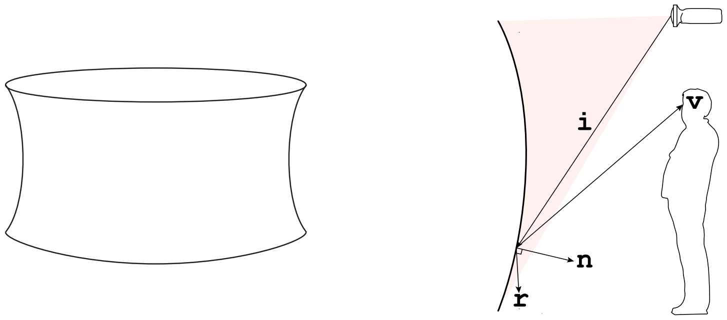


Figure 27 - Catenoidal shape of the stretched screen tends to increase ranges of θ_i, θ_v *and* θ_{vr}. *The degree of deformation is highly exaggerated for purposes of illustration.*

3.5.8 Reducing Angles of Incidence

One possible method for reducing these angles is to use a conical (A. Simon & Göbel, 2002) or bowl-like (Courchesne, 2005; Courchesne et al., 2006; Courchesne, 2007) form, rather than a cylinder. Both these shapes would result in reduced angles of incidence and view, with the curved screen in particular yielding optimal results.

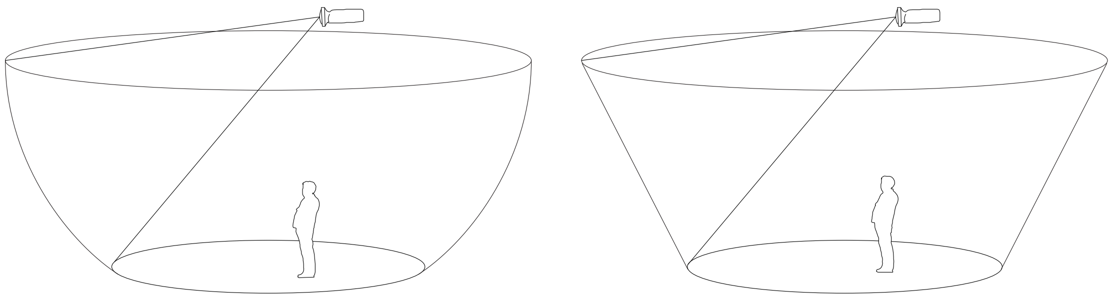


Figure 28 - Conical and bowl-like designs for reducing both the range and variation in θ_{vr}

However, such a curved screen would be difficult to achieve with a soft-fabric screen. Rather, such a solution would require a rigid screen surface, which would be a great impediment to both the portability and the acoustics of the structure. Further, the curved screen introduces a significant non-linear aspect to the distortion of perceived imagery, when viewed by the non-central viewer.

---

[91] The exact shape and extent of distortion depends on the fabric and its behaviour under stress. For a detailed analysis see Bletzinger (1998).

In summary, when considering the geometry of the AVIE and positioning of projectors, careful attention should be given to the angles θ_i and θ_v and in particular the magnitude of difference between θ_v and θ_r. Not only should all these variables be minimised, but to help produce uniformly bright images, the range of these variables over the entire screen surface and throughout the viewing space should be kept to a minimum. In the special case of using polarised light for stereo separation, we must pay extra attention to θ_i, as high angles of incidence may provoke the de-polarisation or re-polarisation of light, leading to stereo 'ghosting' or 'cross-talk.'

3.5.9 Projector Overlap and Blending

In order to achieve a seamless image across all the projectors, the technique of "edge-blending" is adopted. It is a simple concept: the projected images are arranged to partially overlap so that by carefully attenuating the images across the regions of overlap, a single, seamless image of uniform brightness emerges. Even in the rare cases where the projected images take shapes that seamlessly tile without intersection or gaps, blending offers certain advantages. Any discrepancies in colour or brightness between projectors, and any inaccuracies in alignment, will be more visible with non-overlapping 'hard transitions' than with the overlapped, blended 'soft transitions.' In any case, it is often not possible to arrange the projected images without some amount of overlap, which makes some form of blending mandatory.

The early multi-projector panoramas, such as Chase's Cyclorama or Grimoin-Sanson's Cinéorama, achieved a smooth blend between each projected image by carefully positioning masks in front of each projector or on the lens itself. Today this method has been largely rendered redundant by "software-blending," where the images themselves are attenuated prior to being delivered to the projectors. It should be noted that physical blending with a mask enjoys one significant advantage over software-blending, in that it can smoothly blend between projectors with elevated black levels, such as LCD and DLP projectors.

Now, as they lead to an overall reduction in resolution and brightness, blending regions should be as small as possible, while still being sufficiently large to hide the transition from one projector to another. As mentioned above, this is not just a matter of having sufficient blend area for fading from one projector to another imperceptibly, but of minimising the abruptness of change in image characteristics, such as brightness, colour and contrast, which can differ significantly between projectors.[92]

3.5.10 Blending and the Door

Access to the cylinder is provided by an entrance, which should be as narrow as possible, yet still allow passage of a scissor lift for maintenance (80 cm). A swinging door can be used to complete the cylinder, but it is important to note the impact that such a door has on the projector configuration. An open doorway reduces the number of blend regions by one, as no blending is required between the two projectors on either side of the doorway. That is, with an open doorway and n projectors, only $n - 1$ blend regions are required, while a complete 360° screen requires n blend regions for n projectors.

The concept of pixel-efficiency, introduced above, can be extended to reflect this reduction in resolution and brightness caused by blending, by noting that only 50% of pixels within blend regions contribute to the total brightness and resolution.

3.5.11 Shadows

Care must be given that the screen frame, projection rigging or the projectors themselves do not physically intrude in each other's frustums and cast shadows. In addition, attention must be given to the "shadow-free zone," the volume defined by the intersection of all the projector frustums and the ground, which delimits the area within which the audience cannot possibly cast shadows on the screen. As a viewer approaches the screen, the point where they begin to obscure the light from the projector marks the threshold of the shadow-free zone.

---

[92] The source of variation amongst otherwise identical projectors of the same model include differences in optical coatings and the spectral distribution of the lamps, which often changes over time. For an analysis of non-uniform parameters projectors see Majumder and Stevens (2004)

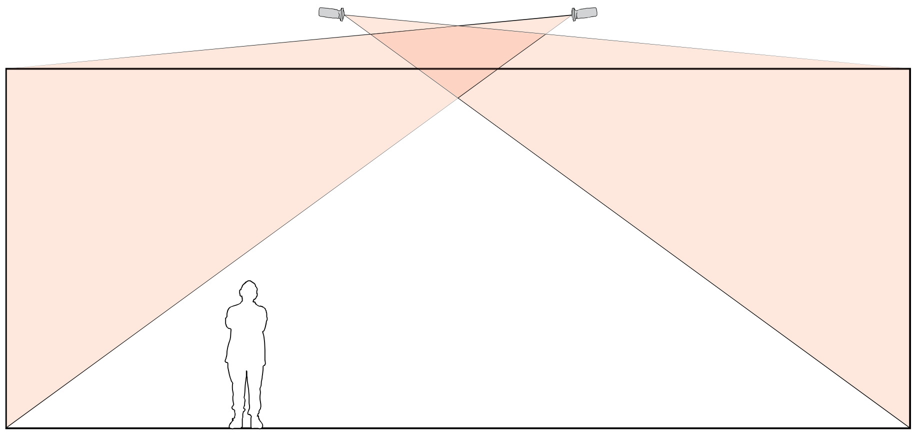


Figure 29 -The shadow-free zone, within which viewers can roam freely without fear of casting shadows on the screen.

3.6 STEREOSCOPIC PROJECTION

Stereopsis is the perception of depth afforded by the simultaneous perception of the world from two different points of view. Stereoscopy is the use of stereopsis to imbue a two-dimensional image with the illusion of depth. A stereoscopic image is, therefore, always composed of two images, one representing the view from the left eye and one from the right eye. The fundamental challenge when creating a stereoscopic display is not, however, the creation of these two monocular images (one need only film or photograph with two cameras, placed side-by-side, or render a computer-generated scene twice, rather than once), but rather the *delivery* of these unique images to the left and right eye. In many respects, the history of stereoscopy is a history of different solutions to the problem of *channel separation* – the delivery of distinct left and right images to the left and right eyes.[93] This section concerns how such channel separation might be achieved in the AVIE.

The simplest, and oldest, stereoscopic displays achieve channel separation by means of some opto-mechanical contrivance – such as mirrors, as in Charles Wheatstone's original 1838 stereoscope, or lenses, in the case of head-mounted displays – to view two physically displaced images. The challenge arises from the geometry of the problem: not only must each image be visible to one eye only, but they must be presented in a manner that still permits binocular *fusion* - the fusing of the two distinct

[93] For a comprehensive account of the development of stereoscopic projection see Zone (2007).

images into a single unitary percept, sometimes referred to as the *Cyclopean image* or *singleness of vision* (Julesz, 1971). For fusion to take place, strict conditions on the position, orientation and size of the two displays, relative to the viewer and relative to one another, must be met. [94] These constraints ultimately lead to the physical attachment of the displays to the viewers head, or vice-versa, and as such can be considered “single-user” (e.g. a head-mounted display) or, in the case of the Kaiser panorama which could be considered “multi-user” in a manner, incompatible with the notion of an unconstrained and physically active audience.

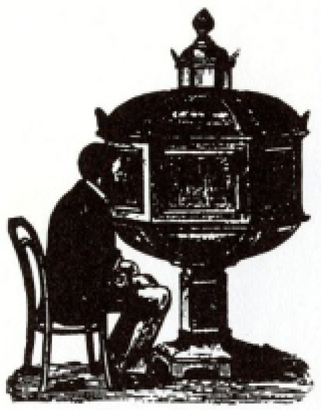


Figure 30 - Kleines Automat Kaiser Panorama, depicted in Gaa and Krüger (1984).

3.6.1 Filtering light

An alternative approach to stereoscopy is the use of a single display for both left and right eye images. The immediate advantage this approach enjoys over the physical displacement method described above is that the two images are guaranteed to occupy the same position, size and orientation relative to the viewer - a vital condition for fusion. The challenge now, however, is displaying two different images in the same physical space, while delivering one image only to each eye.

The solution lies in the transmission of each image in a manner that allows each to be filtered into the appropriate eye on reception.[95] (In engineering terms we might say

[94] David Shafer provides a basic introduction to this geometric problem in his work on an optical device for viewing Salvador Dali’s stereo paintings. See Schafer (1982).

[95] The concept of superposing the two images on a single display and using two different filters to view them can be attributed to Wilhelm Rollman and his 1853 invention of anaglyph stereo (Rollmann, 1853).

that the images are multiplexed upon transmission and de-multiplexed upon reception). This is typically achieved by means of a pair of filters positioned immediately before the eyes, often in the form of spectacles, admitting only the light intended for that eye, and rejecting the rest. Various filtering mechanisms exist and can be classified according to whether they operate by wavelength (anaglyph or spectral comb filtering), by polarity (linear or circular) or time-multiplexing (active stereo), or some combination of these. As will be seen, all of these methods exploit some feature or aspect of light that is ordinarily imperceptible to human vision.

Apart from the various idiosyncratic advantages and disadvantages of each of these methods, which are discussed below, the key performance measures of any stereo separation technique are *light efficiency* and *stereo contrast*. Light efficiency is the percentage of light emitted from the projector's lamp that, after passing through filters during transmission and reception, ultimately reaches the viewer's eye. Stereo contrast is a measure of "cross-talk," "ghosting" or "leakage" arising from imperfections in filtering and transmission.[96] As ghosting increases (and the stereo contrast decreases), the left and right images are no longer perfectly filtered by the viewer's glasses and each eye now perceives a confusing mix of the left and right images. The result is a degradation of image quality and perception of depth, and can rapidly lead to viewer discomfort.[97]

The lower limit for stereo contrast, at which ghosting becomes an impediment to stereo-fusion, depends greatly on the brightness and contrast, stereo baseline and vergence, and the nature of the content of the image being displayed. As such, acceptable lower limits as varied as 300:1 to 20:1 can be found in literature.[98] A commonly used working figure for a lower limit is 100:1.

---

[96] Sometimes also referred to as *extinction ratio*. See Woods (2010, 2012) for a thorough account of the different causes of cross-talk in stereo displays and equally thorough review of the different measures of cross-talk.

[97] See Pastoor (1995); K. Huang et al. (2003); Kooi and Toet (2004); Seuntiëns et al. (2005); Pala et al. (2007); Ukai and Howarth (2008).

[98] See, for example, Yeh and Silverstein (1990); Kooi and Toet (2004); Shestak et al. (2012).

All these methods employ some method of filtering light, which involves the rejection of some proportion of the light (be it by absorption or reflection), and transmission of the remainder. Light is filtered not once, but twice (upon leaving the projector and arrival at the viewer), and as the rejected light is lost forever, light efficiency presents a fundamental concern for all these methods of stereo imaging. Also common among these techniques is an inverse relationship between light efficiency and stereo contrast - increasing one is typically achieved at the expense of the other.

3.6.2 Stereo by polarisation

With stereo by polarisation, the polarity of light is used to distinguish the left and right eye channels, thereby taking advantage of the human eye's blindness to this particular dimension of light. The left and right images are polarised orthogonally to one another, either within the projector or upon leaving the projector, and then transmitted or extinguished upon reception by means of a pair of matching polarisation filters, typically worn as glasses.

The use of polarised light to achieve stereo projection in this manner can be attributed first to John Anderton, who in 1891 used the method to project still stereoscopic images with two magic lanterns (Anderton, 1895). However, it was not until the invention of lightweight plastic Polaroid filters by Edwin H. Land and their application to stereo projection in the 1930s that the technique could be considered viable (Land, 1937, 1940, 1942a, 1942b).

When choosing a polarising filter, several key factors must be considered, including method of polarisation (linear or circular), wavelength range, size, acceptance angle, heat resistance, cost, transmission efficiency, and stereo contrast ratio. Ideally, two perpendicular filters would transmit no light, while the two parallel filters would transmit the theoretical maximum of 50%. In practice, however, this is not the case. Reproduced in Table 2 below are the transmission ratios for the HN-series of linear Polaroid filters - until recently the mainstay of polarising filters - in order to demonstrate the inverse relationship between stereo contrast and light efficiency. These figures vary with wavelength, which accounts for the difference between the nominal and minimum stereo contrast ratio.

Filter	1 sheet Unpolarised -> Polarised Transmission	1 sheet Polarised -> Polarised Transmission	2x Parallel Transmission *The seeing eye* **Light Efficiency**	2x Perpendicular Transmission *The obscured eye*		**Stereo Contrast Ratio**	
				Nominal	Maximum	Nominal	Minimum
HN42	42%	81%	34%	0.5%	0.9%	68:1	38:1
HN38	38%	76%	29%	0.05%	0.1%	580:1	290:1
HN32	32%	63%	20%	0.005%	0.01%	4000:1	2000:1
HN22	22%	45%	10%	0.0005%	0.001%	20000:1	10000:1

Table 2 - Light efficiency and Stereo Contrast for Polaroid Filters.[99]

More recent polarising materials, such as those used in the ITOS XP40HT filter, improve on these figures, achieving up to 32% light efficiency while sustaining a 5000:1 stereo contrast ratio.[100]

Circular versus Linear Polarisation

Light can be polarised either linearly or circularly, the key difference being that linearly polarised light is subject to Malus' Law (Malus, 1809), which describes how the amount of light transmitted through a linear filter changes as a function of the polar orientation of the incoming light, with respect to the filter.

$$I = I_i \cos^2(\theta)$$

where I_i is the incoming intensity,
I the intensity of transmitted light and
θ is the difference between angle of polarisation of the incoming light and the and the orientation of the filter.

Eq. 5 Malus' Law

The consequence of this is that the amount of linearly polarised light transmitted or obscured by a filter, and therefore the amount of cross-talk, is dependent on the tilt of the viewer's head. Were the viewer to tilt their head 45°, each eye would see an equal

[99] Source: *www.knightoptical.com/php/showCatPage.php?cat=103*

[100] ITOS XP40HT accessed September 2013 at *www.itos.de/dateien/polarizer/XP40HT.pdf*

mix of the left and right images and the stereo effect would be utterly destroyed, while at 90° the images would be swapped completely. According to *Malus' Law*, stereo-contrast follows $\cot^2(\theta)$,[101] which falls rapidly from 3000:1 at 1° to 130:1 at 5°, 32:1 at 10° and just 7.5:1 at 20°. To keep within the accepted level of cross-talk, the viewer must tilt their head to the side no more than 5° at any time.

At first this might appear as a compelling reason to abandon linear polarisation in favour of circular polarisation, but unfortunately there are also cases to be heard against circular polarisation. First, for small head-tilt angles, for broad spectrum light, linear polarisers enjoy far greater transmission/extinction values than circular polarisers, due to the quarter-wave retarder component of a circular polariser being wavelength dependent (Sharp et al., 2013). Manufacturers of circular polarisers tend to state transmission/extinction ratios comparable with that of linear polarisers, but these values are for specific wavelengths only and do not accurately describe the passage of broad-spectrum light. In the figures reproduced below, the variation in transmission and obscuration with wavelength for typical linear and circular filters can be easily discerned.[102] Second, linear polarity is preserved better upon reflection than circular polarity (Hong et al., 2010; X. Zhang et al., 2012). Third, circular filters tend to be significantly more expensive than linear filters.

[101] Stereo contrast = transmitted light/obscured light = $I_i \cos^2(\theta) / I_i \cos^2(\pi/2 - \theta) = \cot^2(\theta)$.

[102] ITOS XP40HT. Accessed September 2013 at *www.itos.de/dateien/polarizer/XP40HT.pdf*

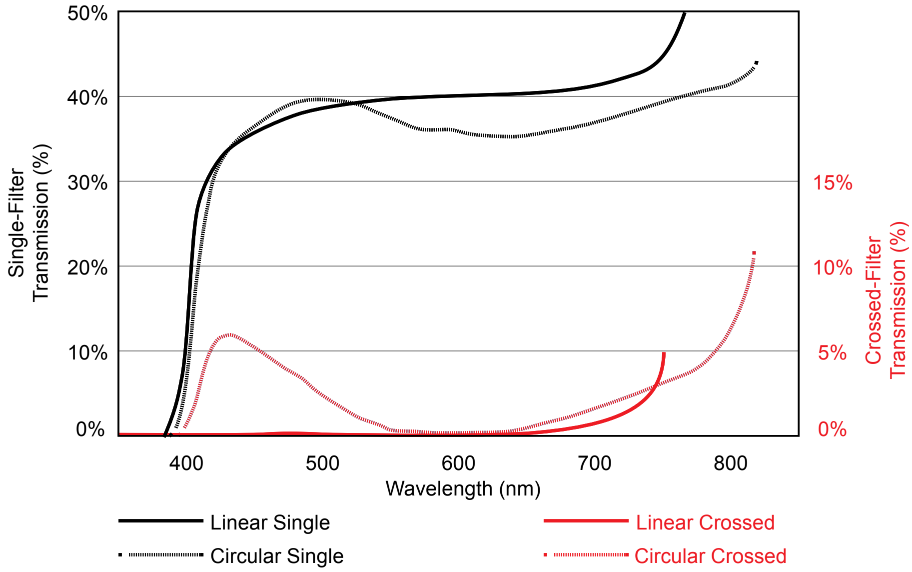


Figure 31 - Transmission vs wavelength for linear and circular filters. Here the transmissiuon characteristics for the linear ITOS XP40HT and circular Polaroid HNCP37 filter are compared. Black lines show transmission for a single filter, while red lines show light leakage for crossed filters. Note the variation in transmission and leakage for the crossed circular filter.

In addition to this, the deterioration in stereo contrast with head-tilt inherent in linear polar light is largely rendered inconsequential by a different problem which emerges with the tilt of the head. For stereo fusion to take place, the images must be displaced along the inter-ocular axis; an axis which shifts with the head. At a tilt of 90 degrees for example, the left and right eyes of the viewer are no longer displaced horizontally, but vertically, and this must be taken into account when creating the images. This necessitates tracking the viewers head, and using this data to generate the images the correct stereo disparity, which ultimately means that any system that encourages significant tilting of the head must be a single-user system, unless all members of the audience can be coerced into tilting their heads in unison. Consequently, the freedom of orientation offered by circular polarisers is not something that can be employed in a multi-user system.

Single or Dual Projector

Stereo by polarisation can be achieved using a pair of projectors, with each projector devoted to either the left or right eye, or with a single projector equipped with a filtering mechanism capable of alternating polarity sufficiently rapidly.[103] By exploiting persistence of vision, the single-projector approach greatly simplifies the system, but at the expense of vastly diminished light efficiency. For while a dynamic filter, such as the *ZScreen (Lipton, 2012)*, exhibits similar transmission ratios to their static cousins (~35%), light is now shared between left and right eyes, instantly halving light efficiency. In addition, as the switch in polarity is not instantaneous, a period within each frame-cycle during which neither eye receives light is required, reducing efficiency even further. The end result is a light efficiency of just 12% (Barco, 2013).

Projector Type

An additional consideration is the type of projectors employed. While DLP[104] and CRT[105] projectors emit unpolarised light, LCD[106] projectors produce light that is polarised linearly during the process of image formation. Although an additional filtering step is still required,[107] if configured correctly, LCD projectors can yield higher light efficiency than their non-polarised counterparts. Woods (2001) documents 57% light efficiency with an LCD projector that produced polarised light consistently across all colour channels.[108]

Angular Sensitivity

A particular weakness of stereo by polarisation is the high degree to which stereo contrast depends on the passage light undertakes before reaching the eye. As

---

[103] For example, a rotating circular polariser with left and right sections, or RealD "ZScreen"; an electro-optical liquid crystal circular polarising filter that alternates polarity with electrical charge. The polarity is typically altered at 120HZ and therefore imperceptible.

[104] DLP: Digital Light Processing - Digital Micro-mirror Device projector as patented by *Texas Instruments.*

[105] CRT: Cathode Ray Tube

[106] LCD: Liquid Crystal Display. Light is polarised as a side-effect as it passes through the LCD image panel.

[107] Unfortunately, often the different colour channels (R, G, B) are not polarised in the same plane, and linear polarisers are required to provide them with a coherent angle of polarisation. And of course the two projectors must have opposite polarity, in which case further filtering is required. For details see Woods (2001).

[108] Experiments were conducted using HN38 polarisers.

mentioned previously, light typically suffers some degree of depolarisation or re-polarisation upon reflection. The degree to which the polarity is changed is a function of both the optical properties of the screen material, as well as the angles of incidence and view (as they were defined in the previous section). High-gain "silvered" screen materials can mitigate the amount of depolarisation, but do so at the expense of uniformity of brightness and range of acceptable viewing angles. Further, the efficacy of the polarising filters changes with the angle of incidence with which light strikes the filter. These angles vary greatly at the projector (the wider the lens, the greater the range of angles), and even more so at the viewer's glasses. The effects of all these angles accumulate and can very quickly reduce stereo contrast below any acceptable level. Circular polarisation is more sensitive to these angular variations than linear. For example, Sharp et al. (2013) measured on a circular polarised system with a wide-angle lens a drop in stereo contrast from 100:1 in the centre of the screen to just 3:1 in the corners.[109]

3.6.3 Stereo by time-filtering: Frame sequential stereo and active-shutter systems

A single projector projects the left and right images sequentially, alternating rapidly between the left point-of-view and the right point-of-view. In concert, a pair of shutters placed immediately before the eyes, open and close in exact synchronisation, temporarily obscuring the left eye when the right image is shown, and vice-versa. This method takes advantage of persistence of vision[110] and the limited temporal acuity of human vision, such that, at sufficiently high frequencies, the intermittent black frames are imperceptible.

Frame-sequential or "active" stereo can be traced back to at least 1897, when Charles Jenkins filed a US Patent for his "device for obtaining stereoscopic effects in

[109] Screen ratio: 1.85:1, lens throw: 0.95, screen gain: 22° half-gain angle.

[110] Persistence of vision is often mistakenly used to explain the perception of motion in a series of still images. While it does explain a flickering image or light source appearing constant above a certain threshold frequency, perception of sequence of static images as a continuous motion is a far more complex phenomenon that cannot be explained solely by persistence of vision.

exhibiting pictures" which employed "a binocular eyepiece" and "a shutter adapted to alternately obstruct the lines of sight through said eyepiece" (Jenkins, 1898). I have found no record of Jenkin's device ever being exhibited publicly, but a similar system - Laurens Hammond's *Teleview* system - was successfully used in 1922 to display one of the very earliest commercial stereo movies (Hammond, 1922).[111] Audience members viewed the film through a rapidly rotating mechanical shutter. Perhaps ahead of its time – for *Teleview* played for just 24 days - active frame-alternating stereo would not resurface until the 1970's when unwieldy mechanical shutters were replaced by opto-electronic liquid-crystal shutters (Lipton, 2012). Today liquid crystal shutters prevail as the enabling technology for active stereo.

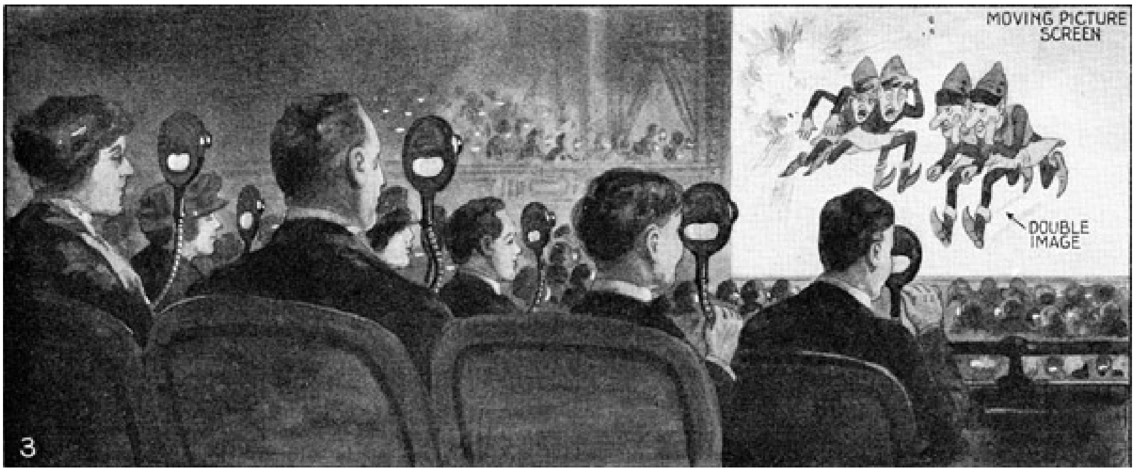


Figure 32 - Teleview active-stereo shutter system. Used in New York, 1922 to display one of the world's first commercial stereoscopic films. (Source: Wikimedia Commons).

The principle advantage enjoyed by active-stereo over polarised stereo is the removal of all constraints and dependencies on angles and orientations of screens, projectors and viewers. Any screen surface can be used, enabling the use of Lambertian screen materials for better uniformity of image.

The disadvantages are unfortunately not insignificant. First, light efficiency is very poor. Theoretically, 50% of the time one eye is obscured, giving a starting efficiency of 50%. However, just as in the single-projector polar solution described above, the shutter

[111] The film, titled *The Man From M.A.R.S.*, was the third ever 3D commercial feature film when it premiered on 27th December 1922. The first two stereo films, both using anaglyphic 3D, premiered also in 1922: *The Power of Love* on 27th September 1922 and *Movies of the Future* just three days prior, on 24th December 1922. For details please see Symmes (2006).

requires time to open and close, and so to avoid a blend of the two images being temporarily visible, a 'blanking period' during which neither eye receives light must be introduced, reducing the light emitted by the projector to around 45%. Further, the glasses themselves are essentially dynamic polarizing filters, and in their open state transmit just 35% of light, resulting in an overall light efficiency of just 13% to 16% (Barco, 2013).

High stereo contrast is possible, up to 500:1, but this is attained at the expense of light efficiency. Increasing the duration of the 'blanking period' produces higher stereo contrast, but reduces brightness, and vice-versa. In practice, stereo contrast is a function of the operational 'duty' cycle of the projector, which varies with different models and technologies, and the speed with which the shutter glasses can switch state.

The shutter-glasses must open and close in perfect synchronisation with the projector, switching between the right and left eyes exactly at the moment the imagery switches from the left to the right point of view.[112] In a multi-projector system such as the AVIE, this implies that all the projectors must operate in perfect synchronisation. Such synchronisation demands the use of hardware 'genlock' synchronisation, a feature only available in the most expensive of graphics cards. This increases the price of the computing infrastructure significantly. The glasses themselves are more expensive, fragile and require batteries, making this system less robust for public exhibition. On the other hand, active stereo requires the installation and alignment of half as many projectors as would be required for a passive solution.

Finally, it should be noted that initial statement regarding a complete independence on angles is not strictly true. LCD active-shutter glasses, when in their open state, are polarising filters. Even if the projector emits unpolarised light, this light will become, to a greater or lesser degree, polarised upon reflection at the screen, leading not just to diminished brightness, but introducing once again a dependency on angles of incidence

[112] This is achieved typically by an infra-red signal or, more recently, by interleaving a special burst of white light into the video signal during the period when both eyes are obscured (DLP-Link).

and reflection, and a non-uniformity of brightness. This will be most significant when using high-gain screens.

3.6.4 Stereo by wavelength

Where previous methods operated on the temporal and polar dimensions of light, a third approach is to filter according to wavelength. With origins as early as 1853,[113] the technique involves encoding the left and right images with different spectra of light, which are then viewed through a pair of wavelength-sensitive filters - that is, coloured glasses. The inherent problem with this method is that while previous methods employed some feature of light imperceptible to humans - polarity or rapid flickering - human vision is very sensitive to changes in wavelength, for this is how we perceive colour. Wavelength based filtering methods, the oldest and most recognisable of filtering solutions for stereoscopy, have therefore always suffered from poor colour reproduction.

A solution here lies once again in the phenomenon of metamerism. As mentioned previously, colour metamerism is the perception of two colours as identical despite being stimulated by light composed of completely different wavelengths. Metamerism provides the basis for *wavelength multiplex visualization*, or *spectral comb filtering*. Developed by Daimler AG in 1999, and marketed under the name *Infitec*, this method capitalises on metamerism to encode the left and right images with slightly different red, green and blue wavelengths, allowing the left and right images to be filtered by glasses and yet be perceived as having the same colours.[114]

Spectral comb filtering appears to encapsulate the best features of both polar and temporal stereo. It is can be used with any screen material and is completely free of any dependencies or constraints on angles of incidence or reflection.[115] Stereo contrast is

---

[113] Wilhelm Rollmann in 1853 invents anaglyph stereo - the bi-chromatic spectral filtering most easily identified by the iconic red and blue glasses (Rollmann, 1853).

[114] For detailed descriptions of the method see Jorke and Fritz (2003, 2006); Jorke et al. (2008); Arnold Simon and Jorke (2011); Jorke and Simon (2012) and Richards and Gomes (2011).

[115] The efficacy of transmission and filtration is most likely, however, very sensitive to the angle with which light passes through the filter.

extremely high, in the order of 1000:1 or even 10000:1 (Jorke & Fritz, 2006). The glasses require no batteries and there are no strict requirements on synchronisation of multiple projectors. And while the metamerism is not perfect, for a colour difference between the two images is perceptible, these differences can be minimised with colour correction algorithms.

The fundamental problem facing this technique is, yet again, light efficiency. The filters extinguish all but a very narrow selection of wavelengths, while the rest is discarded. Barco quotes light efficiencies of 17% for two 3-chip DLP projectors, but as low as 7% for two 1-chip projectors, or a single projector with a rapidly alternating filter (Barco, 2013). However, it should be noted that this technology is relatively new and under active research. Recent results have reported higher 21.5% for 3-chip and 9.6% for 1-chip DLP projectors (Jorke & Simon, 2012). The possibility of using laser illumination to produce the various spectra directly rather than filtering white light, and the creation of filters with more "teeth" in the spectral comb, provide clear avenues for improvement and would suggest that this method is the most promising for the future.

3.6.5 Summary of Stereo Methods

In this section three different methods for achieving channel-separation - the delivery of unique left and right images to the left and right eyes of the viewers - have been introduced. All three methods are viable for use in the AVIE. However, as has been seen, all methods suffer shortcomings, be it in the efficient use of light, cross-talk between the eyes, or sensitivity to angles of reflection or transmission. Table 3 below provides a summary of the comparative strengths and weaknesses when compared to one another.

Filter mechanism	# Projectors	Projector Type	Glasses	Extinction Ratio	Light Efficiency	Strengths and Weaknesses
Linear Polariser	2	CRT or DLP	Passive	Highly variable 300:1 to 1:1[R]	32% - 38%[B]	Pro: • Good light efficiency • No genlock required. • Cheap robust glasses Con: • Variable extinction ratio. • Sensitive to head tilt. • Sensitive to angles. • Requires special screen.
Linear Polariser	2	LCD type 1 R,G,B unaligned	Passive	Highly variable 300:1 to 1:1	32%[W]	
Linear Polariser	2	LCD type 2 R, G, B aligned	Passive	Highly variable 300:1 to 1:1	57%[W] - 59%[B]	
Active Circular Polariser (ZScreen)	1	CRT or DLP	Passive	Poor and highly variable with angle. 80:1 - 3:1[S]	12%[B]	Pro: • Insensitive to head tilt. • Single projector only. Con: • Poor Extinction ratio. • Poor light efficiency. • Highly sensitive to angles. • Requires special screen. • Genlock sync required. • Temporal asynchronicity between eyes.
Liquid Crystal Shutter System	1	CRT or DLP	Active	Good 500:1 - 100:1	12% - 16%[B]	Pro: • Insensitive to head tilt. • Single projector only. • Less sensitive to angles. • Arbitrary screen material. • Good extinction ratio Con: • Poor light efficiency. • Genlock sync required. • Expensive and fragile battery powered glasses. • Temporal asynchronicity between eyes.
Infitec	2	3-chip DLP	Passive	Excellent 1000:1 - 10000:1[J2006]	17%[B] - 21.5%[J2012]	Pro: • Insensitive to head tilt. • Less sensitive to angles. • Arbitrary screen material. • Excellent extinction ratio • No genlock required.

						Con: • Poor light efficiency. • Expensive projectors.
Infitec	2	1-chip DLP or LCD	Passive	1000:1 - 10000:1[J2006]	7%[B] - 9.6%[J2012]	Pro: • Insensitive to head tilt. • Less sensitive to angles. • Arbitrary screen material. • Excellent extinction ratio • No genlock required. Con: • Extremely poor light efficiency.
Active Infitec	1	3-chip DLP	Passive	1000:1 - 10000:1[J2006]	7%[B]	Pro: • Insensitive to head tilt. • Less sensitive to angles. • Arbitrary screen material. • Excellent extinction ratio Con: • Extremely poor light efficiency. • Temporal asynchronicity between eyes. • Genlock required.

Table 3 - Comparison of projected stereo image channel separation technologies.
Key: ***B*** *(Barco, 2013),* ***J2006*** *(Jorke & Fritz, 2006),* ***J2012*** *(Jorke & Simon, 2012),*
R *(Read),* ***S*** *(Sharp et al., 2013),* ***W*** *(Woods, 2001).*

3.7 A QUESTION OF BALANCE

Presented in the previous sections were the different geometric and radiometric factors demanding consideration when designing a cylindrical projection system. The problem can be summarised thus: determine the number of projectors, their position and orientation as well as their throw and aspect-ratio, brightness, contrast, black-levels, lens shift, depth-of-focus and resolution, and the radius, height and gain of the cylindrical screen, that result in maximal aggregate resolution, brightness, uniformity and sharpness of image, maximal range of the shadow free zone and maximal quality of the stereo separation, while minimising image distortion and cost.

As these factors enjoy a complex inter-dependence, where increasing one may reduce another, what follows is a process of optimisation and constraint satisfaction. For example, choosing wide-angle (short-throw) lenses reduces the number of projectors needed and increases the shadow-free zone, but increases the depth of focus required, reduces pixel-efficiency and overall brightness, increases distortion, and, because of an increase in the range of angles of incidence and angles of view, leads to a reduction of uniformity of both brightness and polarity. Long-throw lenses increase the number of projectors, increasing resolution and decreasing angles of incidence, but greatly increases cost and reduces the shadow-free-zone. These are just some of the interdependencies that must be considered.

In order to balance all these interdependent factors and help arrive at the optimal type, number and arrangement of projectors, 3D simulation was performed by Paul Bourke using custom software.[116] In addition a simple parametric simulator, called *AVIEator*, was developed to rapidly explore the effect of varying different parameters.[117]

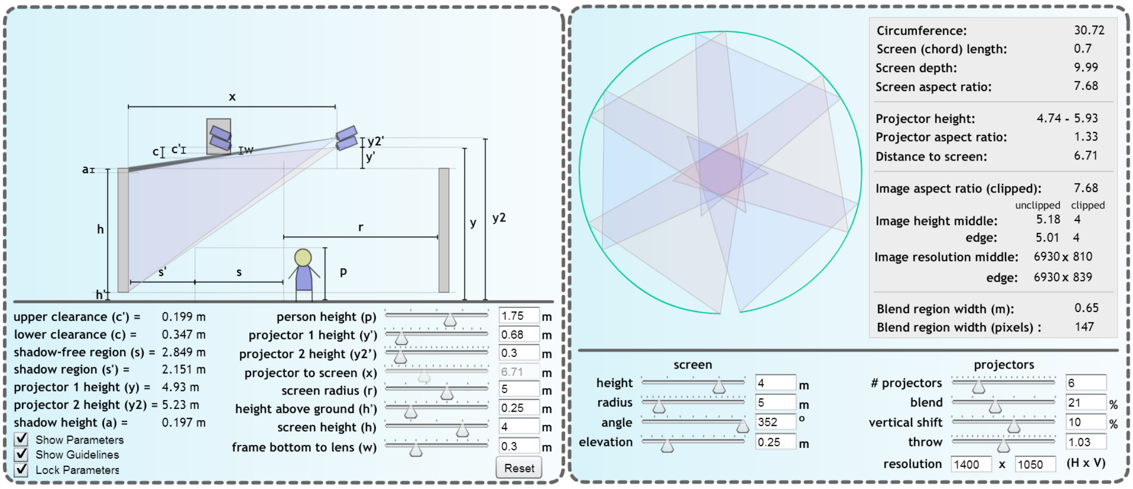


Figure 33 - AVIEator configuration software (screenshot).

3.8 SOLUTION

At the time of writing, twelve complete AVIE systems have been constructed, each with slightly different parameters. Almost all the differences between the systems however, fall into three categories: the dimensions of the cylinder; the method of stereo channel separation; and the presence or absence of a closing door.

---

116 See *paulbourke.net/miscellaneous/cylmapper/index.html*.

117 *AVIEator* was developed by iCinema Engineer Alex Kupstov under the author's supervision.

The AVIE screen takes the form of a cylinder, 10 metres in diameter and 3.985 metres in height.[118] A narrow 80 cm opening provides entry. A door, made from the same material as the projection screen, can be attached here to the screen frame and swung closed to complete the 360° image, if desired.[119]

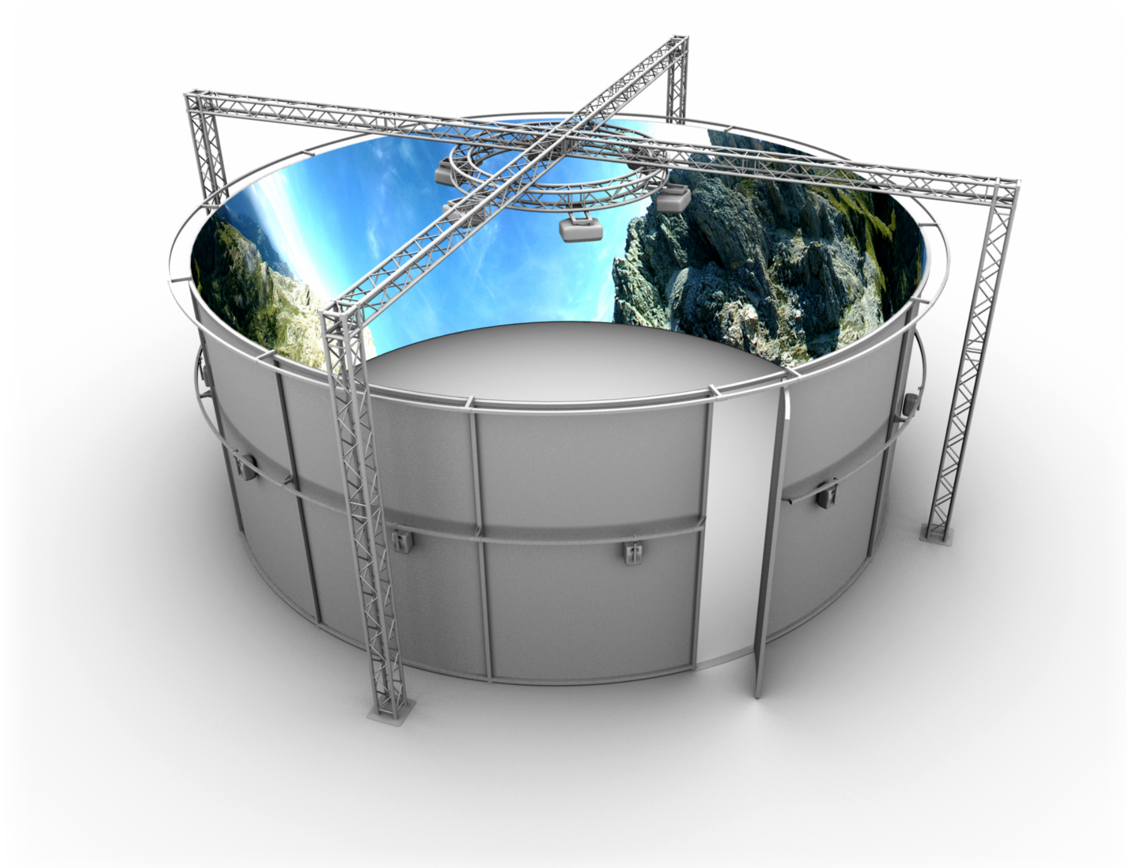


Figure 34 - The AVIE system.

The screen is illuminated by six projectors (in the case of active stereo), or six pairs of projectors (in the case of passive stereo), each with an image aspect ratio of 4:3, a resolution of 1400x1050 and with wide-angle lens providing a 1:1 throw ratio.[120] The

---

118 The first system built, was 10m in diameter and 3.6m high, due to the limited height of the ceiling. Other systems have had slightly smaller diameters due to room size. The most common configuration, however, is 10 meters diameter, 3.985 metres high. In all systems the doorway is between 80 and 85 cm wide.

119 The original system had no swinging door. With some systems, the doorway is only 2.1m high, above which the screen continues uninterrupted, while in others it extends the whole height of the cylinder, as pictured above.

120 More recent installations of the AVIE use just 5 projectors, each 1920x1080 with a 16:9 aspect ratio, producing a similar resolution overall.

projectors are arranged in a circle approximately 2.5 metres in diameter, projecting inward across the centre of the cylinder to the opposite side of the screen, so as to illuminate the entire cylinder. As the projectors shoot across the centre of the circle, and therefore across each other, they must be sufficiently raised above the top of the screen to avoid casting shadows of each other. In turn, this demands that the projectors be slightly tilted downward, just to the point where the bottom of the cylinder is completely illuminated, as they do not have sufficient vertical lens shift to reach the bottom without tilting. Each projector illuminates a 66.4° segment of the cylinder.

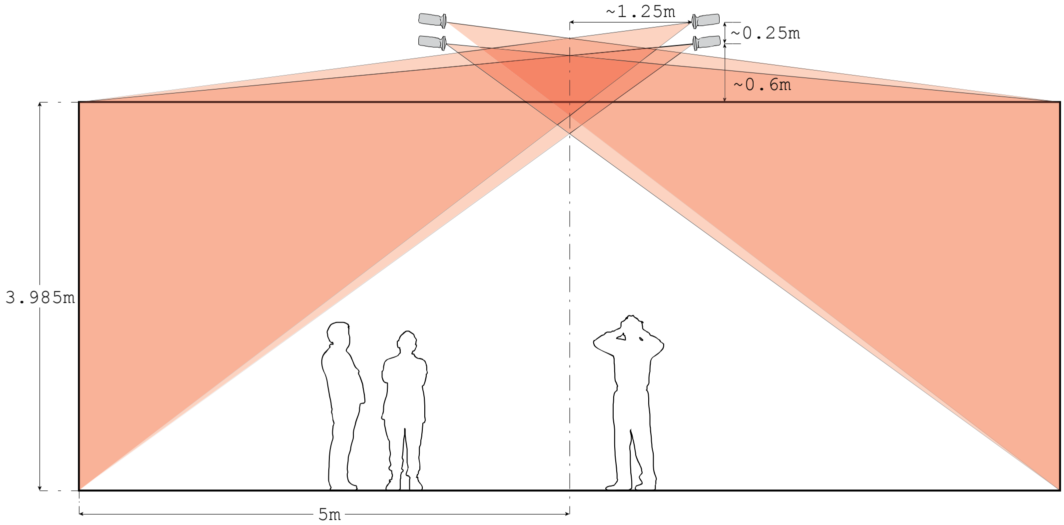


Figure 35 - AVIE dimensions.

In the initial design, six pairs of projectors with linearly-polarising filters were used to achieve channel separation. Linearly polarising filters were chosen over circular-polarising variety because of the higher and more uniform stereo contrast ratios. Later incarnations of the AVIE adopted an active stereo solution with viewers wearing LCD shutter-glasses, and as such requiring only 6 projectors. In the early stages of development, tests with *Infitec* spectral comb filters were performed but colour distortion and poor light efficiency were considered too greater impediments to offset the associated advantages.

There is a slight difference between those AVIE systems with a closing door and those without. Those without a close-able door, need only illuminate approximately 351°, as the doorway occupies the remaining 9° and no blending is required between the two projectors on either side of the doorway. Those with a close-able door must light the full 360°, plus an extra blending region. However the pixel-density for both the 351° and 360° configurations are the same, as the distance from the screen to the projectors is limited not by horizontal field-of-view of each projector but by the vertical field-of-view. The only difference, therefore, between the 360° and 351° configurations is the number and size of the blend regions. With 6 blend regions and 360°, each blend region is 6.4°, or 9.6% of the projectors width. With 5 blend regions and 351°, each region is 8°-9° in width, or 12%-13% of the projectors width.

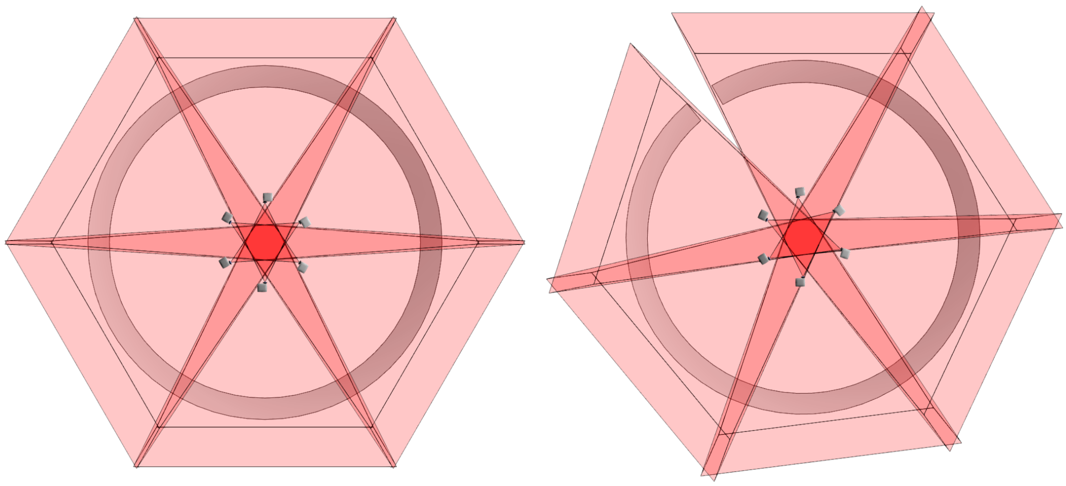


Figure 36 - AVIE Projector Overlap.
Left: Projector arrangement for full 360° projection with closing door.
Right: Arrangement for 351° projection and permanently open door. Note only 5-blend regions.

Resolution

This configuration results in a panoramic display approximately 7570 pixels in circumference,[121] and 900-1000 pixels in height. Pixel density is non-uniform; the vertical resolution varies from approximately 1035 at the left and right edges of the each projectors image, to just 900 in the middle. 90% of pixels emitted by the projectors strike the screen.

Note that for the 12-projector passive system, there is a minor difference in vertical resolution and pixel efficiency between the top ring and bottom ring of projectors, due to the 30cm displacement between the top and bottom rings. The top projectors enjoy a vertical resolution ranging from 875 to 1010 pixels, and 88% pixels strike the screen.

[121] 7570 around the full 360°, 7400 for the 351° configuration.

3.9 Omnistereo Rendering and Projection

This section concerns the projection, both mathematically and physically, of a virtual light field onto a cylindrical image surface, so that this image might serve as a visual surrogate of the light field for multiple viewers. It begins with the *omnistereo projection,* a cylindrical stereo mapping that minimises (without completely removing) view distortion when multiple viewers are free to turn and move about within the cylinder. Methods for rendering real-time omnistereo images of virtual polygonal worlds are discussed, after which discussion turns to the physical projection of panoramic images onto a cylindrical screen using an array of projectors. This part concerns the correction of projection distortion and the blending of projected images to achieve a seamless, undistorted panoramic image.

3.9.1 Introduction

To recap: a projective image is formed by tracing the intersection of rays of light with an image surface. When those rays of light intersect at a single point, they define the *centre-of-projection* of the image, and when the image is viewed from this point, the image may serve as a (geometrically) faithful surrogate of the light-field. As human vision is binocular, the light field must be represented by two images captured from the points-of-view of the left and right eyes. It is useful to consider these two images as a single stereo image captured from a single point-of-view (the point between the two eyes), and direction (defined by the stereo-baseline - the orientation and length of the ray between the two eyes).

A stereo image, when viewed from any point or direction other than that with which it was created, will appear distorted. As the position of the viewer deviates from the point-of-view of the image, the world appears to compress, dilate or shear. Similarly, as the direction of viewer deviates from the direction captured within the image, discrepancies in the orientation and magnitude of binocular disparity give rise to distorted perception of depth. With a 360° image, it is these discrepancies in binocular disparity that are most disruptive, for as the viewer turns, they perceive an image in which stereo-disparity is in places completely extinguished or reversed, giving rise to highly conflicting depth cues and, possibly, a failure of stereo fusion, when the viewer is no longer able to perceive the images as a single stereo image.

The omnistereo projection seeks to minimise distortion for all viewers, by first assuming the viewer to be located at the centre of the cylinder, and second, by abandoning the notion of a single direction of view, and assuming the direction of view to be everywhere perpendicular to the screen surface. This latter assumption produces a stereo panorama with an equal binocular disparity in all viewing directions. The result is a panoramic image that provides a perceptually undistorted view for all viewers, but does so by sacrificing geometric accuracy in the periphery of vision.

3.9.2 Related Work

Early examples of omnistereo projections can be found in mobile robotics, where omnistereo vision has been applied to the problem of localisation and mapping. In 1989, Sarachik (1989) equipped a robot with two cameras, each treated as vertical-slit cameras, by sampling the images from a single vertical line only. As the robot turned pirouettes, features points could be tracked as they pass across the two vertical slits, endowing the robot with a rudimentary form of omnistereoscopic vision. It was also in the context of mobile robotics that Ishiguro et al. (1990, 1992) first described in detail the capture of a stereo panorama with a single rotating camera. Their system involved rotating a camera around an axis displaced from the camera's centre of projection, and sampling the images along two vertical slits. The robot, by identifying the same object as it passes through the two slits, could derive the distance to the object from the camera's rotation, and construct a 3D map of its environment. The authors demonstrate, as a by-product of their vision system, the construction of omnistereo panorama by stitching together the vertical lines of pixels obtained from the two slits over the 360°. Zheng and Tsuji (1990, 1992) also equip a mobile robot with a single slit-camera so that it might construct panoramic views of its environment. They generalize the construction of 'panoramic views' for a slit-camera moving on circular, linear or arbitrarily curved paths through space, and provide analytical expressions of the distortion a horizontal line undergoes when projected into these different panoramic views.

The use of stereo panoramas for image display begins with H. C. Huang et al. (1996); H. C. Huang and Hung (1998), who use a pair of rotating video cameras to generate the panoramas. They do not however treat the cameras as vertical-slit cameras, but instead employ a very elaborate system of image registration, warping and equalization, stitching and blending to form a seamless panorama from conventional planar images. The resulting stereo panorama is viewed on a regular monitor, and the mouse is used to navigate the image.

In 1998, Naemura et al. (1998) used the omnistereo projection to display images in a CAVE VR system, thereby providing an "adequate depth sensation" for multiple viewers regardless of their direction of view. Named "multi-user immersive stereo", it is omnistereo in all but name. They do not, however, display real-time imagery but an omnistereo projection of a static 4D light-field captured with rotating camera (Naemura et al., 1997).

With a series of publication beginning in 1997,[122] Peleg, Ben-Ezra and Pritch describe the use of a conventional (2D) video camera as a 1-D slit sensor, allowing a moving camera to capture arbitrary shaped "manifold projections," multi-centre-of-projection images formed by stitching lines captured from a moving camera. They demonstrate the use of a single rotating camera to form an omnistereo panorama and show how the stereo-base line can be dynamically adjusted throughout the panorama to better accommodate the local depth in the scene, by simply moving the slits left and right in the sampled images. In 2000 (Peleg et al., 2000), they outlined two designs for an omnistereo camera without moving parts, one employing a "spiral mirror" and the other a "spiral lens," although it does not appear that such lens has ever been built. A summary of their diverse work on omnistereo panorama can be found in Peleg et al. (2001), in which they also outline the rendering of omnistereo panorama

[122] (Peleg & Herman, 1997; Peleg & Ben-Ezra, 1999; Peleg et al., 2000; Pritch et al., 2000; Peleg et al., 2001)

from 3D virtual models by approximating the non-linear omnistereo projection with a patchwork of planar projections.

Inspired by Peleg's work on manifold mosaics, Shum et al. (1999) discuss the capture of dual-strip panoramas by rotating a single perspective camera before extending the concept to "concentric mosaics," a 3D formulation of a light field as a set of concentric omnistereo panoramic images. The basic insight is that by rotating a video camera off-axis through 360°, and retaining the whole video-frame at each point, sufficient information is captured to reconstruct a stereo panorama *from any point* within the circle circumscribed by the camera's motion (Shum & He, 1999). They demonstrate the real-time rendering of novel views, which were presumably viewed on a conventional monitor (Shum et al., 2000).

Pajdla (2002) generalises the concept of an image projection (or 'camera'), showing that all cameras fall somewhere between the 'central camera,' where all rays intersect at a single point, and the 'oblique camera,' where no rays intersect at all. In between these two extremes, one finds 'pushbroom cameras,' where all rays intersect on a line and the omnistereo panorama, where all rays intersect on a circle. These 'cameras' describe the image formed by moving a camera along these various paths. Similarly, Seitz and Kim (2002) analyse the space of all possible projections and ask, of all the possible mappings, which possess horizontal parallax and may therefore be viewed as stereoscopic images? By generalising the notion of epipolar geometry to fit multi-perspective (non-central) projections, they find that only those projections that are described by a camera moving along a *conic curve* (point, line, ellipse, parabola or hyperbola) give rise to stereo images:

Projection / Camera	**Camera Path**	**Epipolar surface**
Perspective	point	pencil of planes
Omnistereo (stereo panorama)	circle	half-hyperboloid
360°x360°	circle	hyperboloid
Spherical omnivergent	sphere	pencil of planes
Pushbroom stereo	line	pencil of planes
Stereo cyclograph	ellipse	half-hyperboloid
Parabolic panorama	parabola	hyperbolic paraboloid

Table 4 - Space of all stereo projections. Reproduced from Seitz and Kim (2002).

According to Seitz and Kim, this result is surprising for two reasons: a) that "we can potentially fuse images that have multiple centres of projection" and b) that "so few varieties of stereo views exist – out of all possible 2D subsets of the 5D set of rays, only three varieties satisfy the stereo constraint" (2002).

Shimamura et al. (2000a, 2000b) constructed a cylindrical projection theatre for the display of stereo panorama. They describe the display of panoramic video with depth maps computed from vertical disparity, and the mixing of CG elements, but it is not clear if they employed an omnistereo projection. Explicit use of the omnistereo projection for the display of 3D models is, however, given in Andreas Simon et al. (2004), where they demonstrate two methods for the real-time omnistereo rendering of a 3D model. The first is the "multi-view" method in which the omnistereo projection is approximated by a number of conventional single-point planar projections.[123] The second is "object-warping," which appears to involve the distortion of a model's vertices to simulate the non-linear omnistereo projection. These methods were used to display real-time 3D content in a three-sided CAVE and 240° cylindrical display.

In 2004, in the context of the iCinema Centre project *Conversations*, the author developed a system for the viewing of high-resolution spherical omnistereo video through head-mounted displays. These videos were created by compositing live action 'green screen' footage onto spherical omnistereo photographs. The resulting videos were viewed through stereo head-mounted displays (McQuire & Papastergiadis, 2006).

Couture et al. (2010) offer an analysis of the error introduced by the omnistereo assumption for a centrally located viewer, finding it to be largely below the acuity of stereo perception. More recently, Couture and Roy (2013) have described a method for capturing omnistereo video using a triad or quintet of fish-eye cameras arranged in a circle, but rather than pointing outwards, they are all oriented upwards. They show that such an arrangement of cameras allows the reconstruction of an omnistereo panorama for the point at the centre of these cameras, albeit with some distortions.

[123] The 240° panorama is divided into 15° strips, each rendered with a standard planar-perspective projection.

3.9.3 Omnistereo Projection

The essence of the omnistereo projection is elegantly revealed by writing it as a mapping from the plenoptic function L to an image I. While a traditional single-point monocular panorama, captured from point of view **p**, may be written as:[124]

$$I(u, v) = L(p_x, p_y, p_z, \theta, \phi)$$

where I is a panoramic image,
L the plenoptic function,
(p_x, p_y, p_z) the point of view,
and u, v are image-space coordinates
$0 \le u \le 1$, $0 \le v \le 1$

Eq. 6 Monocular Panoramic Projection

with

$$u = \frac{\theta}{2\pi} \qquad v = \frac{R \tan \phi}{H} + \frac{1}{2}$$

where H, R are the cylinder height and radius,
$0 \le u \le 1$, $0 \le v \le 1$, and
$0 \le \theta \le 2\pi$ and $-H/2R \le \tan(\phi) \le H/2R$

Eq. 7 Image-space to Polar coordinates

For no reason other than the sake of brevity, we now re-parameterise the image function I to directly accept polar-coordinates over range $\theta \in [0, 2\pi]$ and $\phi \in [\tan^{-1}(-H/2R), \tan^{-1}(H/2R)]$.[125] So if a monocular panorama is described by $I(\theta, \phi) = L(p_x, p_y, p_z, \theta, \phi)$, then an omnistereo panorama may be written as:

$$I_L(\theta, \phi) = L(p_x - b\cos\theta, p_y, p_z + b\sin\theta, \theta, \phi)$$

$$I_R(\theta, \phi) = L(p_x + b\cos\theta, p_y, p_z - b\sin\theta, \theta, \phi)$$

Where b is half the stereo base-line

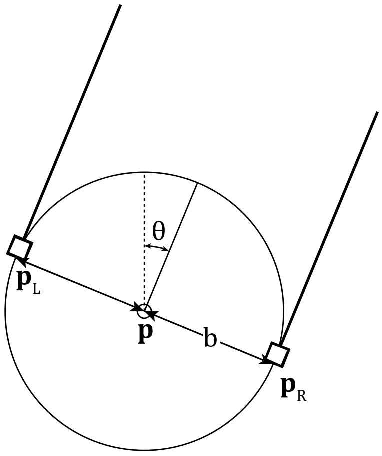
Eq. 8 Stereo Panoramic Projection

As can be seen, where the monocular panorama has single-point **p** as centre-of-projection, the centre-of-projection for the stereo panorama is a circle with radius b, centred on **p**. For simplicity, we set $\mathbf{p} = 0$, and omit it from the equation.

This configuration can be conceived as a pair of virtual vertical-slit cameras with positions $\mathbf{p}_L = (-b\cos\theta, 0, b\sin\theta)$ and $\mathbf{p}_R = (b\cos\theta, 0, -b\sin\theta)$:

[124] In the following equations, u and v serve as coordinates into the cylindrical image-space, with u being the horizontal axis (normalised from 0 to 1 around the circumference), and v the vertical axis (normalised 0 to 1 along the height of the cylinder).

[125] This introduces a non-linearity in the vertical dimension of the image-function, but this can be ignored for in the discussion here.



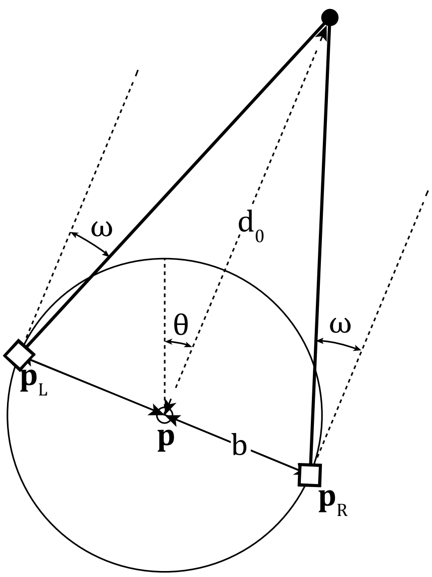
$$I_L(\theta, \phi) = L(\mathbf{p}_L, \theta, \phi)$$

$$I_R(\theta, \phi) = L(\mathbf{p}_R, \theta, \phi)$$

Figure 37 - Parallel slit-cameras.

Zero-Parallax and the Horopter

The term "horopter" refers to the locus of points in 3D space which, for a given stereo configuration, have zero binocular disparity, or zero parallax. In other words, the horopter is the set of points in 3D space that are projected onto the same image coordinates for both the left and right eye. For the omnistereo projection, the horopter is the circle circumscribed by the intersection of rays from the left and right view-points. In the formulation above, the rays are parallel, and so the radius of zero-parallax is at infinity. The horopter can, however, be set arbitrarily by rotating the virtual cameras inwards by an equal amount ω so that their rays intersect at a desired zero-parallax distance d_0. There are, however, two ways this might be achieved. In the first, the cameras are rotated about their own axis:



$$\omega = \tan^{-1}\left(\frac{b}{d_0}\right)$$

Where d_0 is the desired zero-parallax distance, b the half stereo base-line and ω the necessary rotation of camera.

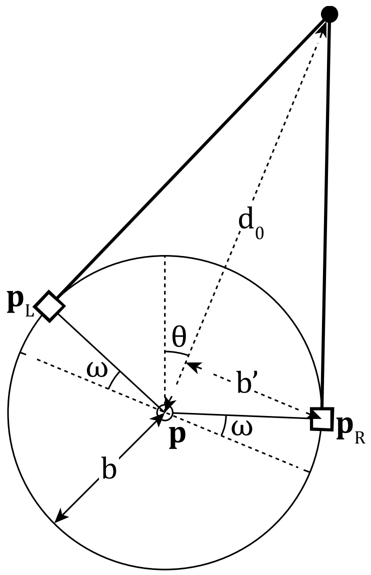
This yields the mapping:

$$I_L(\theta, \phi) = L(-b\cos\theta, 0, b\sin\theta, \theta + \omega, \phi)$$

$$I_R(\theta, \phi) = L(b\cos\theta, 0, -b\sin\theta, \theta - \omega, \phi)$$

Figure 38 - Controlling the depth of zero-parallax with locally rotated camera.

In the second, the cameras are rotated about the central viewpoint **p**, thereby changing the position of $\mathbf{p}_L$ and $\mathbf{p}_R$:



$$\omega = \sin^{-1}\left(\frac{b}{d_0}\right)$$

Where d_0 is the desired zero-parallax distance, b the half stereo base-line and ω the necessary rotation of camera.

Note the reduced stereo base-line b' = b cos(ω).

This yields the mapping:

$$\mathrm{I_L}(\theta\,,\phi) = \mathrm{L}(-\mathrm{b}\cos(\theta+\omega)\,,0,\mathrm{b}\sin(\theta+\omega)\,,\theta+\omega,\phi)$$

$$\mathrm{I_R}(\theta\,,\phi) = \mathrm{L}(\mathrm{b}\cos(\theta-\omega)\,,0,-\mathrm{b}\sin(\theta-\omega)\,,\theta-\omega,\phi)$$

Figure 39 - Controlling the depth of zero-parallax with centrally rotated camera.

The two projections are not equivalent, for they sample the plenoptic function over a different set of coordinates. Note that the second method results in a diminished effective stereo base-line b' = bcos(ω). In the AVIE, the horopter is typically set at the screen surface (d_0 = 5), and with a stereo baseline of b = 0.065, the reduced base-line b' = 0.064999. In other words, the reduction is negligible.

Rotating the panorama

When displaying the images, it is possible to rotate one image with respect to the other. This can be effected by adding or subtracting an offset (z) to the u-coordinate when sampling the image for display I(u + z, v):

$$\mathrm{I_L}(\theta+\mathrm{z},\phi) = \mathrm{L}(-\mathrm{b}\cos(\theta+\mathrm{z})\,,0,\mathrm{b}\sin(\theta+\mathrm{z})\,,\theta+\mathrm{z},\phi)$$

$$\mathrm{I}_R(\theta-\mathrm{z},\phi) = \mathrm{L}(-\mathrm{b}\cos(\theta-\mathrm{z})\,,0,\mathrm{b}\sin(\theta-\mathrm{z})\,,\theta-\mathrm{z},\phi)$$

Eq. 9 Rotating the panoramas in image-space

These equations reveal that rotating the images has precisely the same effect as rotating the cameras around the central axis, as depicted in Figure 39 above, and is therefore a simple and effective means of controlling the horopter in both pre-rendered and real-time imagery.

3.9.4 Rendering Omnistereo

Both the mono and stereo panoramic projections are non-linear, meaning that straight lines in the 3D world are transformed into curves in image-space. With the mono panorama, a horizontal line in 3D space, with distance d from the origin and height h, traces in image-space the sinusoid:

$$v = \frac{R\,h\cos(2\pi u)}{H\,d} + \frac{1}{2}$$

Where h is the height of the horizontal line, and d its distance from the origin. H, R are the cylinder height and radius.

Eq. 10 Mono panorama distortion of a horizontal line

With the omnistereo panorama, the horizontal line is described by:[126]

$$v = \frac{R\,h\cos(2\pi u)}{H\,(d - b\cos(2\pi u))} + \frac{1}{2}$$

Where h is the height of the line, d the distance from the origin and b is half the stereo base-line. H, R are the cylinder height and radius.

Eq. 11 Omnistereo distortion of a horizontal line

The primary challenge faced when implementing a real-time implementation of the omnistereo projection is that contemporary real-time graphics is very much reliant on the acceleration afforded by graphics-processing units (GPU) to achieve interactive real-time frame rates, and these devices are specifically designed to accelerate linear, single-point perspective projections. The omnistereo projection is neither linear nor single point, so in order to take advantage of these powerful computing devices, the omnistereo projection must be approximated in a manner amenable for computation on a GPU.

The two methods implemented here are *slice-rendering*, in which the panorama is treated as a sequence of thin vertical linear projections, and *vertex-projection,* in which the vertices of the scene are projected non-linearly, but interpolation (rasterisation) between vertices remains linear.

Slice-Rendering

As outlined by Peleg et al. (2001), the omnistereo projection can be approximated by a series of single-point planar-perspective projections. We approximate the cylindrical screen with a faceted cylinder made up of equal sized planar faces, and for each these

[126] Zheng and Tsuji (1990).

faces, render a conventional stereo image with two cameras, displaced to the left and right of the face normal. Note that this requires the use of asymmetric "off-axis" cameras; a commonly used technique when rendering stereo images.

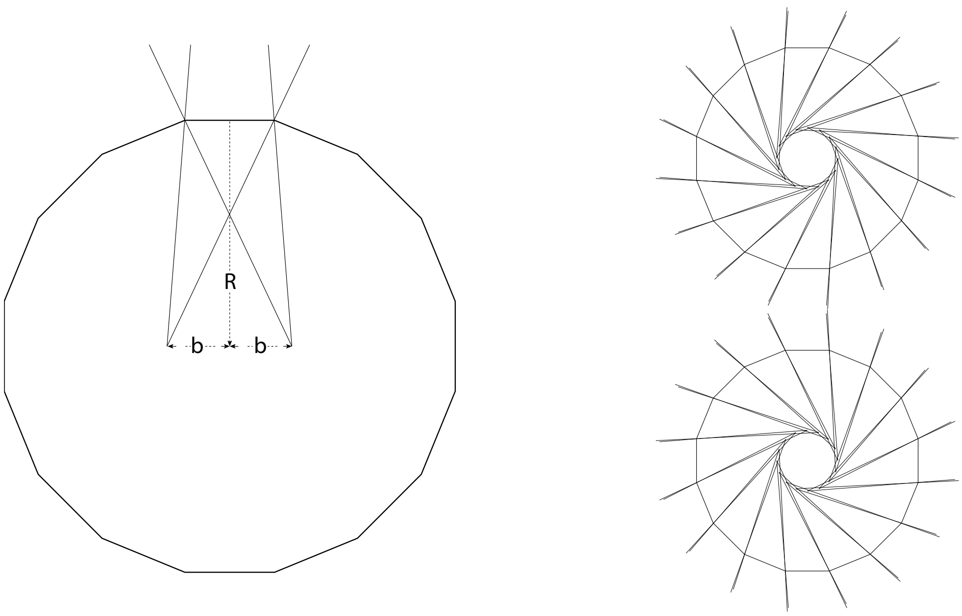


Figure 40 - The Image-slice Omnistereo approximation. Shown here with n=16 slices. Left: a pair of asymmetric cameras, separated by distance 2b, rendering a stereo image for one slice. Right: the left and right panorama are constructed by rotating the cameras around the central axis.

The distortion introduced by this approximation can be understood by examining the distortion of a horizontal line when projected onto a cylinder, as described in *Eq. 10.*[127] A planar projection of a horizontal line, with height h and distance d maps to a straight, horizontal line in image-space

$$v = \frac{R\,h}{H\,d} + \frac{1}{2}$$

Where h is the height of the horizontal line, d its distance from the origin. H, R are the cylinder height and radius.

Eq. 12 Perspective projection of a horizontal line.

The vertical error introduced by the slice-render approximation is then the difference between *Eq. 10* and *Eq. 12*:

$$\Delta v = \frac{R\,h}{H\,d}(1 - \cos\theta)$$

Eq. 13 Maximum vertical disparity between planar and cylindrical projection of a horizontal line.

[127] As b << d, as is the case in the AVIE, the difference between Eq. 10 and Eq. 11 grows inconsequential.

From this equation it is clear that vertical error grows with the height (h) of the line above or below eye level, diminishes with the distance (d) of the line and, most importantly, grows as the projected point moves horizontally away from the centre of the slice (θ), according to 1-cos(θ). Setting R = d = 5 and H = 4 and h = 2, and assuming a vertical resolution of 1000 pixels, then distortions no greater than a single pixel are assured with slices of 7.25°, which in the AVIE, corresponds to roughly 8 slices per projector. In practice, as few as 4 slices per projector, where the 15° slices give rise to deviations of roughly 4 pixels (1.7cm) are adequate. This is, however, predicated on d ≥ R. When virtual objects enter within the perimeter of the AVIE, the error grows.

Another source of error arising from the slice-rendering approximation is found in the small misalignment between neighbouring frusta. As the cameras are rotated around the circle of projection, their frusta slightly overlap and diverge, as is illustrated in Figure 41. The angle between frusta can be expressed as function of stereo-baseline, cylinder radius and the number of slices:

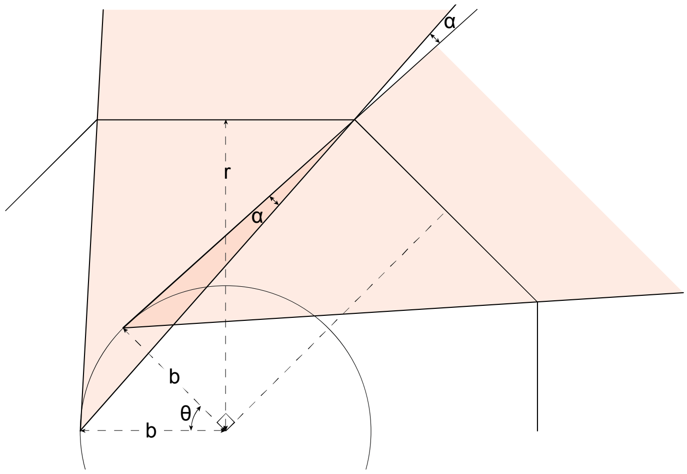


Figure 41 - Omnistereo frustum overlap.

$$\alpha = \tan^{-1}\left(\frac{w\cos\theta + \tan\frac{\theta}{2}}{1 - \mathrm{w}\sin\theta}\right) - \tan^{-1}\left(\mathrm{w} + \tan\frac{\theta}{2}\right)$$

Where $w = \frac{b}{r}$ and $\theta = \frac{2\pi}{n}$

See Appendix F for proof.

Eq. 14 Omnistereo frustum overlap error

However, with a stereo baseline of b = 0.065 / 2, a screen radius of 5 m and 24 slices, the angle of overlap is imperceptible at just 0.0006°. Nonetheless, there is an arrangement of cameras that eliminates entirely this overlapping of frusta. By rotating the cameras around the origin, while keeping the outer faces of their frusta fixed, an arrangement is reached where the gaps between neighbouring frusta are extinguished.

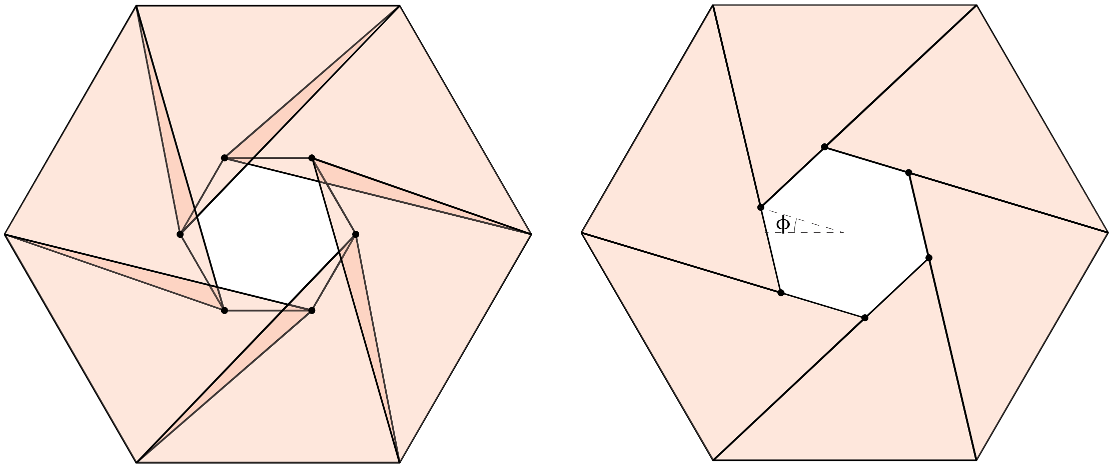


Figure 42 - Seamless arrangement of omnistereo cameras.
Left: Conventional slice-rendering with overlapping frusta.
Right: The cameras are rotated by angle ϕ so that their frusta perfectly align.

The angle ϕ by which the cameras must be rotated is given by:

$$\phi = \sin^{-1}\left(\frac{\left(\mathrm{r} - 2\mathrm{b}\sin\left(\frac{\pi}{\mathrm{n}}\right)\sin\theta - \mathrm{r}\frac{\sin^2\theta}{\cos^2\left(\frac{\pi}{\mathrm{n}}\right)}\right)}{\mathrm{b}}\right)$$

where

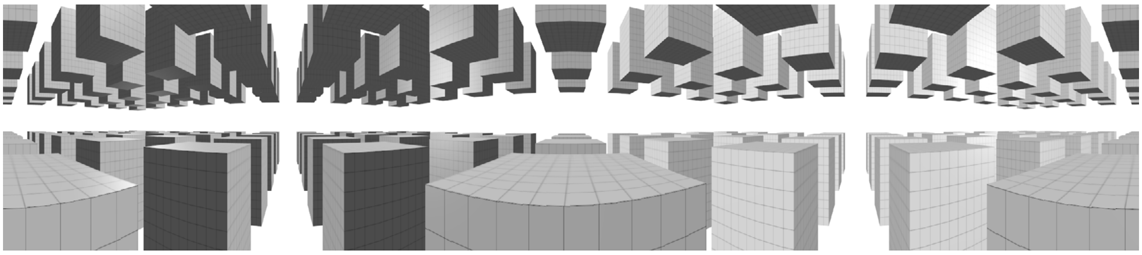
$$\theta = \cos^{-1}\left(\frac{\sqrt{z^2(k^2-1)+2(k+1)-z(k-1)}}{2}\right)$$

with
$z = \frac{b}{r}\cos\left(\frac{\pi}{n}\right)$,
$k = \cos\left(\pi - \frac{2\pi}{n}\right)$,
n - number of slices,
b - stereo baseline and
r - radius of the screen.

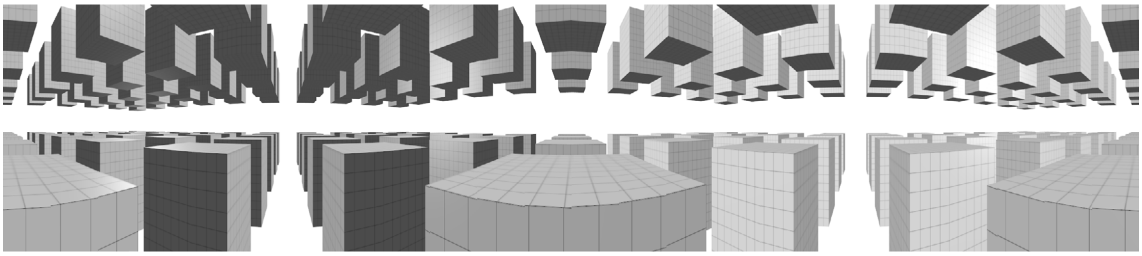
Eq. 15 Omnistereo frusta alignment

Unfortunately, even when frusta are aligned in this manner, vertical discrepancies between neighbouring slices remain, due to neighbouring cameras having different distances to a virtual object lying on the transition between frusta. In fact, as the size of this vertical discrepancy depends on the difference in distance between the virtual object and the neighbouring cameras, the arrangement pictured in Figure 42 and Eq. 15 actually slightly exasperates this problem, as it ensures that the difference in distances between object and camera for objects lying on the transition between frusta is the maximum possible.

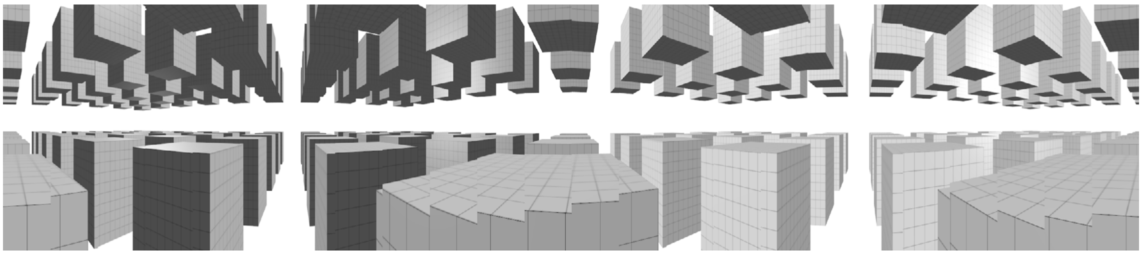
The closer the virtual object to the centre of the cylinder, the greater the vertical discrepancy between slices. Fortunately, with as little as 24 slices (15° each) and an inter-pupil distance of 6.5cm, these discrepancies are imperceptible for all but the closest objects, and only present a visible distraction when virtual objects are within 50 or 100cm of the centre of the AVIE.



c) normal interpupil distance = 6.5cm, slices n = 48.



b) normal interpupil distance = 6.5cm, slices n = 24.

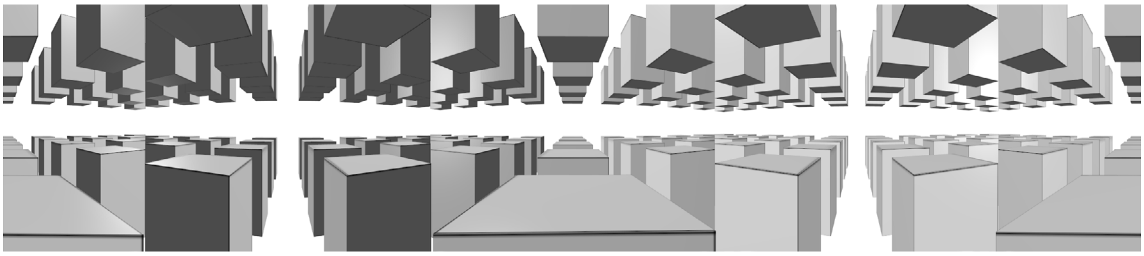


a) highly exaggerated interpupil distance = 50cm, slices n = 24.

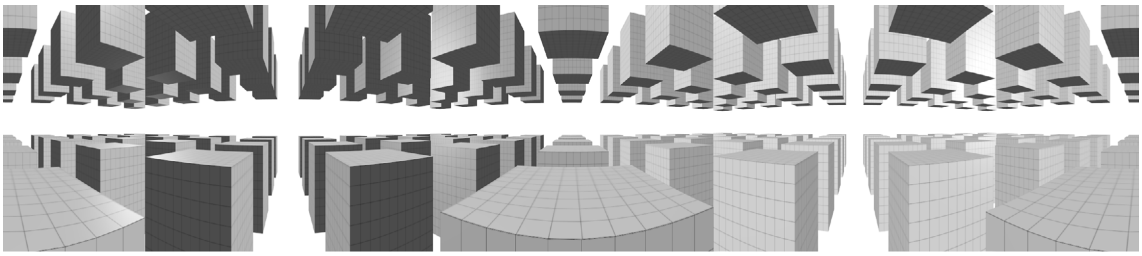
Figure 43 - Vertical discrepancies in omnistereo slice rendering.
A regular grid of cubes, all 1m in size and separated by 1m gaps, is rendered with the slice method. The closest cube face is 50cm from the centre of the cylinder. Note, these images are shown with doubled height to render the discrepancies more visible.

Vertex-Projection

An alternative to the slice-rendering method is to replace the planar perspective projection of vertices with the omnistereo projection, by taking advantage of the programmable "vertex shaders" of today's GPU. This method eliminates the problem of discrepancies between slices altogether. However, interpolation between vertices remains linear, and so suitably tessellated models are required if an undistorted image is to be achieved.[128] Recent GPU offer highly flexible tessellation units capable of dynamically subdividing meshes on-the-fly, suggesting an ideal mechanism for omnistereo rendering. Unfortunately these tessellation units have only been available since the release of the OpenGL 4.0/DirectX11.0 APIs, neither of which are supported by the iCinema SDK/*Virtools* platform (see Section 3.13). For more on this method see Ardouin et al. (2014).



a) With only 1 quad per cube face, the cubes are incorrectly rendered with straight lines.



b) With 10x10 quads per cube face, the cubes are appropriately distorted.

Figure 44 - Omnistereo vertex-projection rendering. Interpolation between vertices remains linear, so models must be highly tessellated.

Both the slice-rendering and vertex-projection techniques are available for use with the iCinema SDK. In the past, however, the slice-rendering method has been favoured for its greater flexibility. The vertex-projection method demands that all models be rendered with a specific vertex-shader, which can complicate the use of 3rd party content, especially if it contains many shaders of its own. In contrast, the slice-rendering method is able to render any 3D content without modification.

[128] Interpolation between vertices, even in the most recent GPU, is hardwired to be linear.

3.9.5 Projection Distortion and Blending

Having devised a method of rendering a virtual 3D scene into a stereo-panorama, the problem of physically projecting this panorama onto a cylindrical screen is now addressed. This involves two problems: correcting distortion introduced by projecting a planar image onto a curved screen, and blending overlapping projected images to create a single seamless image.

Distortion correction is achieved by employing the same principle underlying anamorphosis. Namely, one distortion can be used to counter-act another. By pre-distorting the projected image in a specific manner, distortions introduced during projection can be undone entirely. Blending is achieved by attenuating the image in the overlap regions so that the combined light of the two overlapping images remains constant and equal to the light of just one projector.

To achieve this, during the final stages of image generation, the image to be projected is mapped onto a suitably shaped *distortion mesh*, and then modulated using a suitable *blend texture*. Calibrating the projection system involves calculating a unique distortion mesh and blend texture for each projector.

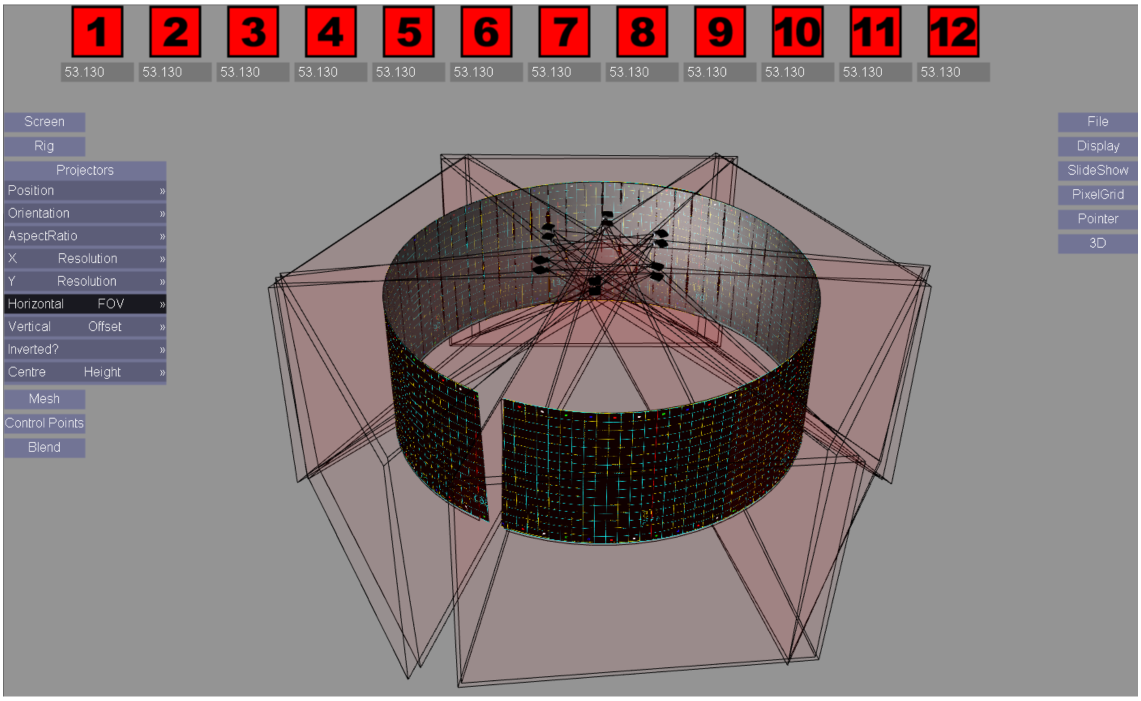


Figure 45 - icAVIEConfig *projector calibration tool. Geometric model.*

To achieve this, a tool named icAVIEConfig was developed to allow interactive calibration. First, real world measurements of the theatre are fed into a simple parametric model that captures the salient features of the theatre. The screen is modelled as a cylinder, with a gap for the door and virtual projectors with variable position, orientation, throw and vertical shift. From this model, first approximations of the distortion meshes are calculated by 'inversely' projecting a regular grid on the cylindrical screen back into the frame of each projector. In other words, the distortion mesh is simply the image of a regular grid on the cylinders surface, as seen from the point of view of the projector.

Second, discrepancies between this basic model and the real-world (for example, the real screen is not a true cylinder, and the virtual projector parameters do not perfectly describe the real projectors), are accommodated by manipulating the distortion mesh with simple horizontal and vertical scaling and translation, linear and parabolic keystone, as well as three separate parabolic curves for vertical displacement of the top, middle and bottom of the mesh.

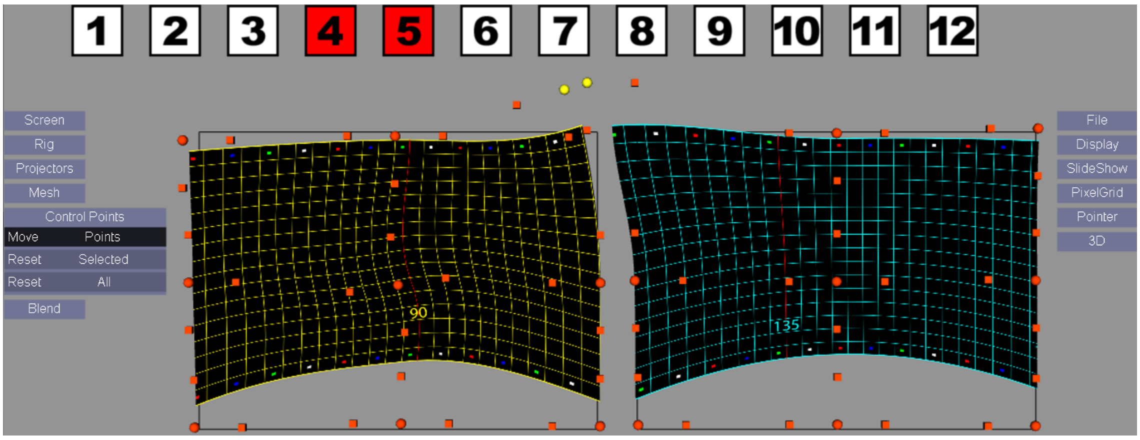


Figure 46 - icAVIEConfig projector calibration tool. Distortion meshes.

Finally, for pixel accurate alignment, the user manipulates a nine-control point quadratic Bezier patch to fine-tune the mesh.

The process of calibration takes place within the AVIE itself. icAVIEConfig projects a calibration grid in the theatre, which, at first, appears distorted and broken. Calibration is complete when this grid appears uniform across the whole projected surface, which is achieved by using the three tools described above.

The blend textures are generated using a simple model that defines the start and end points of each blend region, plus a power function and gamma correction, according to a method described by Bourke (2004). This is slightly modified to allow the start and end point to vary according to a pair of vertical curves. Blending is an interactive process of adjusting the start and end curves and gamma parameters for each overlapping region. With experience, the AVIE can be calibrated completely within an hour or two.

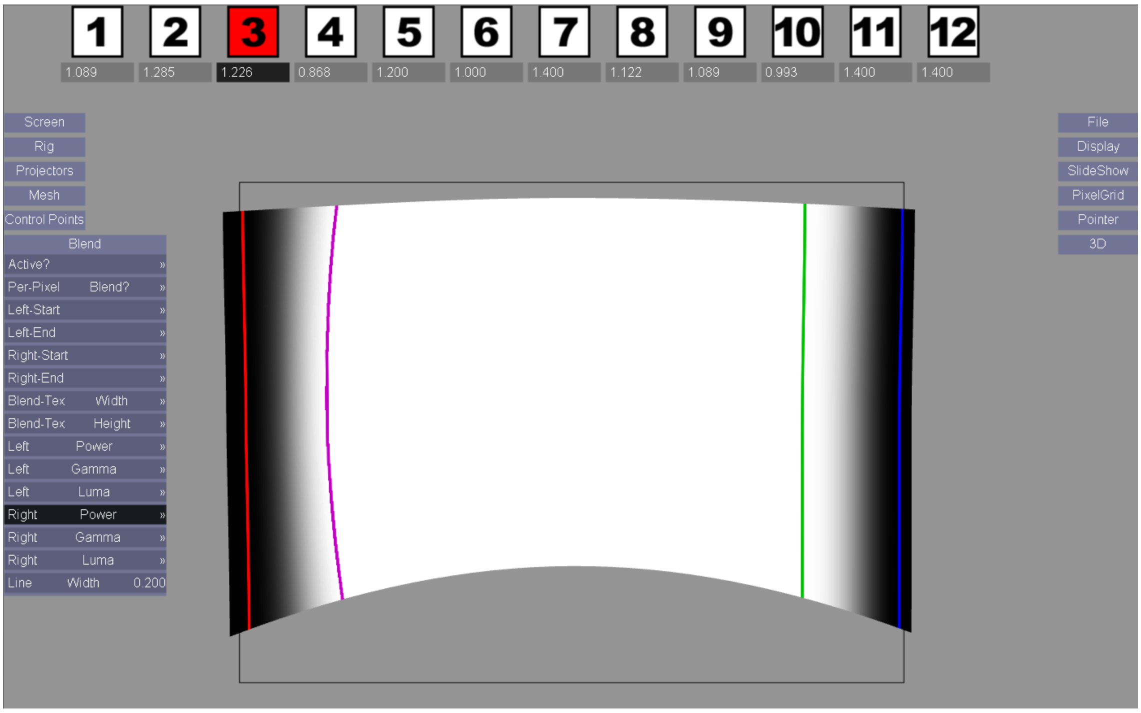


Figure 47 - icAVIEConfig projector calibration tool. Blend textures.

3.10 AUDIO

While the discussion until this point has been concerned mostly with the perception of light, the perception of sound is held to play an equally important role in the construction of presence.[129] The AVIE is, therefore, designed as to be as much a sonic medium as it is a visual one, and to this end the AVIE is furnished with a surround spatial sound system. The term "surround" refers to the distribution of speakers around the audience, and the capacity of the system to envelope the audience within a pervasive 360° field of sound. "Spatial," on the other hand, refers to a more subtle and more powerful feature of the system, which is the ability to produce sounds that are perceived to emanate from arbitrary positions in space. A spatial sound system, therefore, permits the creation of sonic environments with perceptible spatial structure, in which sonic entities have distinct direction, distance and size. The calculation of these spatial soundscapes is dynamic and real-time, allowing not only for arbitrary movement of sounds through space but real-time and interactive movement *through* such soundscapes. And when coupled with the AVIE's projection system, it permits the presentation of virtual worlds rich with spatially coherent visual and sonic cues. These are all the elements necessary for presence: active perception of the environment, ego-motion, and binding of multi-modal cues.

Immersive perception of sound is, in some respects, analogous to the immersive perception of light. In place of the light field we speak of a pervasive ambient *wave field*, which we continually sample over a shifting set of coordinates. One manner to replicate this perceptive relationship with the wave field is the approach dictated by *physical mimesis*, which is to reproduce the wave field in its entirety. Such an approach is known as *holophony* – the sonic equivalent of holography – and implementations of this approach include *wave field synthesis* (Berkhout, 1988; Berkhout et al., 1993; Boone et al., 1995; Spors et al., 2004; Spors et al., 2008) and *higher-order ambisonics* (Poletti, 2000; Daniel, 2003; Daniel et al., 2003). A holophonic sound system is inherently multi-user,

---

[129] For introductions to the importance of sound in the generation of presence, see Gilkey and Weisenberger (1995); Gilkey et al. (1999); Murray et al. (2000).

non-invasive and allows listeners to wander freely within the sound field; all features that nominate holophony as an ideal solution for sound within the AVIE.

However, despite certainly being more tractable than its visual counterpart (for practical large-scale real-time wave field systems currently exist,[130] which cannot be said for holography), the technique was not considered appropriate for the AVIE due to the high number of loudspeakers required. With wave field synthesis, the principal source of error is *spatial aliasing* caused by the gaps between speakers. Rabenstein and Spors propose a distance of 10cm to 20cm (2005) (and elsewhere 30 cm (2006)), between each speaker,[131] suggesting 100 to 300 loudspeakers would be required for a circle the size of the AVIE; a number held to be too high when considered in terms of portability, computability and affordability.[132]

An alternative approach is binaural synthesis, in which the wave field is not reproduced everywhere, but only at the listener's ears - a feat most easily achieved by wearing a pair of headphones. Here, the listener's position and orientation must be tracked in real-time so that their 'point-of-view' in the virtual wave field may be simulated accordingly, using such techniques as *head-related transfer functions* (e.g. Cheng & Wakefield, 1999). This approach is perfectly analogous to the use of a head-mounted display, where also by tracking the position and orientation of the user's head it is possible to simulate only the portion of the light-field that strikes the eyes.

Binaural synthesis was also not considered for the AVIE, somewhat due to the technical difficulty of wirelessly tracking the position and orientation of all the listeners with sufficiently low latency and accuracy, and the amount of computation that would

---

130 A number of functional wave field synthesis systems have been successfully implemented, with the largest to date being the 2700 loudspeaker system (832 channel) at the Technische Universität Berlin (Moldrzyk et al., 2007). See Baalman (2007) for a somewhat recent review of other research prototypes and applications.

131 Spatial aliasing in wave front synthesis is a function of the distance between speakers, frequency and angle of the wave-front. With 10cm between speakers, spatial aliasing occurs at frequencies above 1700Hz, while at 30cm, aliasing begins around 566Hz. See Spors and Rabenstein (2006) for a mathematical analysis of spatial aliasing arising in a circular array.

132 And even with this large number of speakers, sound sources are constrained to a 2D plane. To move virtual sounds vertically, thousands of speakers would be required.

be required to compute the binaural signals for all users and all sound sources, in real-time. Both these challenges might be overcome, but the main reason for not adopting such an approach is that wearing headphones prohibits natural communication between users, which is one of the design demands of the AVIE. For an example of binaural synthesis in which both the problem of tracking and inter-user communication were solved for a three-user system, the reader is directed to the iCinema project *Conversations* (Jin et al., 2005; Kan et al., 2005).

It is possible to achieve binaural spatialisation without headphones, by tracking the position and orientation of a listener, and using techniques such as cross-talk cancellation to 'project' the binaural wave field from loudspeakers (e.g. W. G. Gardner, 1997; Kuhlen et al., 2007). In essence, this is a sound system in which the 'sweet-spot' – the small volume of space within which the wave field is correctly perceived – is extremely narrow, but by dynamically shifting this sweet-spot to follow the position of the listener's ears, the illusion of a continuous perceptual space is conjured. Sometimes referred to as *transaural synthesis*, this method is perfectly analogous to the head-tracked imagery of VR systems like the CAVE (Cruz-Neira et al., 1993). And just like the images in the CAVE, this method only works for a single user at a time. As such, it is also not considered for use in the AVIE.

For a recent synopsis and comparison of these techniques and others, the reader is directed to Blauert and Rabenstein (2012). In addition, see Wiggins (2004) and Lossius (2007) for gentle introductions to the subject of psychoacoustics, spatial sound perception and spatial sound rendering.

3.10.1 Amplitude Panning

The spatialisation technique adopted here is *amplitude panning*. Amplitude panning exploits the psychoacoustic phenomenon of *summing localisation,* where two or more coherent [133] sounds arriving from different directions can, under certain specific conditions, be heard as a single 'phantom' sound emanating from a single position in

[133] i.e. identical wave patterns differing in amplitude or phase only.

space (Warncke, 1941; Blauert, 1997). The phantom sound is perceived as originating from a point somewhere between the two speakers, and the position of this point is influenced by the difference in amplitude and phase of the two sound signals. The apparent position of the phantom sound can therefore be controlled by carefully manipulating the relative amplitudes (amplitude panning) or phase (time-based panning), or both at once, of the source signals. This is indeed the principal of stereophony: sound emanating from two loudspeakers can be perceived as originating from a point somewhere between the two speakers, and the position of this phantom source can be controlled.

Unlike the holophonic methods of wave field synthesis or higher order ambisonics, amplitude panning is *not* an attempt at physical mimesis. That is, spatialisation is not achieved through the reproduction or even approximation of the sound-field that would exist were there really a sound emanating from that point in space. The wave field produced by amplitude panning is quite different from that of the real one, and it is solely due to a peculiar feature of human hearing - summing localisation - that a phantom sound is perceived. In this sense, it is an example of *metameric mimesis*, where very different physical stimuli give rise to indistinguishable perceptions.

Further, the phantom sounds produced by amplitude panning could only be said to be truly indistinguishable from their 'real' counterparts when the listener is within a relatively confined 'sweet-spot.' Outside of this, the perception of space offered by amplitude panning is distorted. It will be shown however that the system tends to fail *gracefully*, such that, despite these distortions, the system provides adequate spatial accuracy to support immersion. In this respect, amplitude panning could be considered the visual analogue of omnistereoscopic projection, which too is only strictly correct for a central viewer, but degrades smoothly as the viewer moves afield. Most importantly, it will be shown that the distortions are never so much that sonic and visual cues result in perceptual conflicts, which would be most destructive to presence. In fact, the opposite is the case: when combined with visual cues, both sound and vision are subject to frequent *cross-modal enhancement*, which is to say that the visual cues enhance the perceived fidelity of the sonic cues and vice versa.

Amplitude panning, when compared to the holophonic methods, is computationally simple, allowing real-time spatialisation of large numbers of sounds with a single ordinary computer. However, with this simplicity comes limitations. First, unlike the holophonic methods in which any manner of sound field might be reproduced, amplitude panning is limited to the reproduction of soundscapes constructed entirely from *point sounds.* Second, amplitude panning only concerns the perceived *direction* of a sound, and does not concern the perceived *distance* to a sound. As such, additional methods are required to provide sounds with the illusion of distance or size, so that sounds may appear to emanate from within or without the AVIE, or appear to emanate not from a point, but from an area or volume in space. Similarly, amplitude panning does not in itself address the simulation of acoustic effects of the environment, such as reverberations, reflections or occlusions. For such effects, additional methods are also required.

Amplitude panning as a spatialisation technique follows directly from the pioneering work of Alan Blumlein in the 1930s (e.g. Blumlein, 1958), and is the basis of common stereophony. According to Blauert and Rabenstein, the method was extended to large arrays of speakers in the 1960s, when it was known as a "synthetic sound field," "and intensively used for scientific purposes, for example, at the Technical University of Dresden and the University of Göttingen" (Blauert & Rabenstein, 2012). They cite Karl-Heinz Stockhausen's Spherical Concert Hall, Germany's contribution to the 1970 World Expo in Osaka, as an example of spatialisation by amplitude-panning, but it is not clear how much the psychophysical phenomenon of summing localisation was intentionally involved in the spatialisation of sound.

In 1975, Theile and Plenge (1977) analysed in detail the spatial accuracy of a six speaker circular system based on amplitude-panning, within which they propose six as a minimum number of speakers for a spatially stable 360° system. This work is referenced in detail below. More recently, it is the work of Ville Pulkki and colleagues at the Helsinki University of Technology that offers the most thorough development and analysis of the method. In a series of articles published from 1997 to 2002,[134] Pulkki et al. generalise amplitude panning to 2D circular and 3D spherical speaker arrays with arbitrary numbers of speakers, provide a computationally efficient "vector-based"[135] method for computing loud speaker volumes, and provide an empirical evaluation and theoretical analysis of the resulting systems. This body of work is the starting point for the analysis of the AVIE sound system presented here.

3.10.2 Sound System Design

The initial AVIE prototype had 24 speakers arranged in two rings of twelve, one at the foot of the screen and the other at the top. This configuration was adopted because the screen of the AVIE was originally fabricated from an unperforated material that was not completely transparent to sound, and it was considered that placing the speakers behind such a screen would be detrimental to sound quality. All subsequent AVIE systems, however, have been constructed with acoustically transparent perforated screens that permit the placement of speakers behind the screen, without diminishing sound quality. This allows the use of just 12 speakers, positioned at head height in a ring behind the screen. In both cases, two sub-woofers are used to provide non-spatial low frequency signals.

The speakers are connected in pairs to six[136] stereo amplifiers, which are fed audio signals by a single multi-channel audio interface - a form of high-end external sound card[137] which permits a single computer to drive the entire speaker array as a single integrated system.

As the 24-speaker system is the exception, unless otherwise stated, the discussion that follows will concern only the 12-speaker system.

---

[134] (Pulkki, 1997, 1999; Pulkki et al., 1999; Pulkki, 2001a, 2001b; Pulkki & Karjalainen, 2001; Pulkki, 2002). For a summary of this work in a single publication, the reader is directed to Pulkki and Karjalainen (2008).

[135] Hence the name *vector-based amplitude panning* or VBAP - a name which has become synonymous with amplitude panning in general.

[136] In the case of the 24-speaker system there would be 12 stereo amplifiers.

[137] In most instantiations of the AVIE, the RME *Fireface 800* is used.

3.10.3 Panning Functions

By panning between neighbouring speakers, a phantom sound can be positioned anywhere on the circle circumscribed by the speaker array. Sounds are bound to this circle, in which case the desired position of the virtual sound can be described by a single angle of azimuth ϕ. The simulation of distance is achieved by other methods, discussed below.

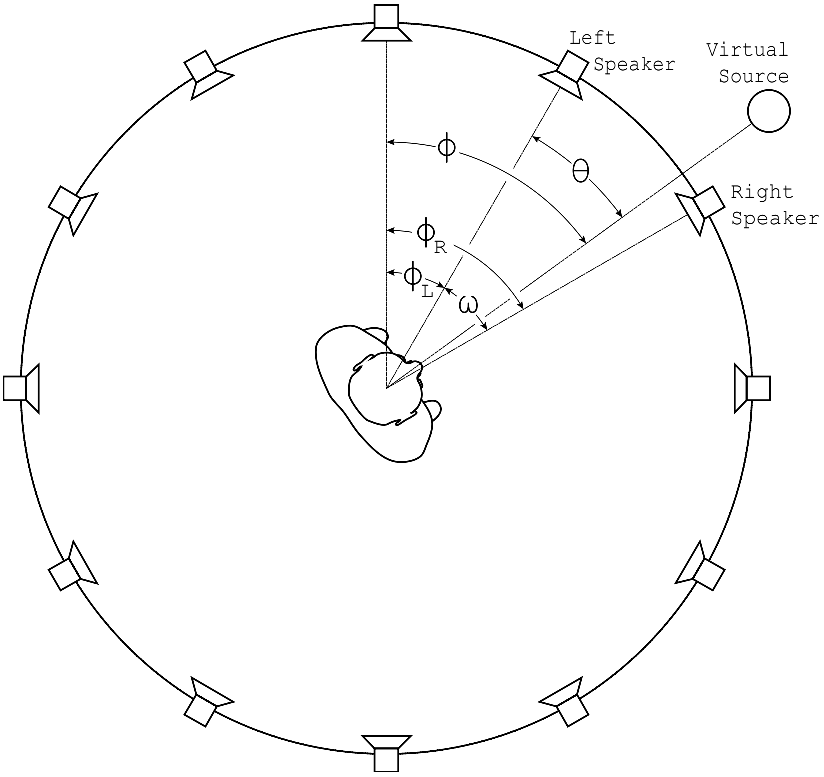


Figure 48 - Amplitude Panning Angles.
ω is the angle between speakers. ϕ the desired azimuth of the phantom sound. ϕ_L and ϕ_R are the azimuth of the left and right speakers, with $\phi_L + \omega = \phi_R$. θ is the local angle of the phantom sound within the closest speaker pair.

The algorithm proceeds like this: given a desired virtual sound position, the azimuth angle ϕ is calculated. This angle is used to select the closest pair of speakers with azimuth angles ϕ_L and ϕ_R. Let ω be the angle between speakers: $\phi_L + \omega = \phi_R$. The gain levels g_L and g_R of the left and right speakers are therefore a function of the local angle $\theta = \phi - \phi_L$. Now, a pair of *panning functions* are needed that, given the local angle θ of the sound position, provide the relative gains of the left and right speakers. These functions, g_L and g_R, should be continuous over domain $0 <= \theta <= \omega$, and symmetrical, such that:

$$g_L(\theta) = g_R(\omega - \theta)$$ *Eq. 16 Panning function symmetry*

They should have range over the unit interval (g_L, g_R $\in$ $[0,\ 1]$), and begin and end at 0 and 1:

$$g_L(0) = g_R(\omega) = 0,$$
$$g_L(\omega) = g_R(0) = 1$$ *Eq. 17 Panning function range*

As a sound is moved around the circle, the apparent volume should remain constant. So, following Pulkki (1997), the combined energy of the two speakers should be constant for all θ:

$$\sqrt{g_L^2(\theta) + g_R^2(\theta)} = 1$$ *Eq. 18 Panning function constancy*

There are any number of functions that satisfy these conditions. Two commonly used pairs of stereo panning functions that satisfy both these conditions are:

$$g_L(\theta) = \cos\left(\frac{\theta\pi}{2\omega}\right)$$
$$g_R(\theta) = \sin\left(\frac{\theta\pi}{2\omega}\right)$$ *Eq. 19 Cos/Sin Panning functions*

and

$$g_L(\theta) = \sqrt{1 - \theta/\omega}$$
$$g_R(\theta) = \sqrt{\theta/\omega}$$ *Eq. 20 Square-root Panning functions*

However, while these panning functions are mathematically compliant, they do not reflect the psychoacoustics of summing localisation. For example, it has been shown that when the listener is oriented towards the midpoint of the two speakers, the perceived position of the phantom sound follows the stereophonic *sine law* (Clark et al., 1957; Bauer, 1962):

$$\frac{\sin(\omega/2 - \theta)}{\sin(\omega/2)} = \frac{g_L(\theta) - g_R(\theta)}{g_L(\theta) + g_R(\theta)}$$ *Eq. 21 Stereophonic Sine Law*

This *sine law* is satisfied by

$$g_L(\theta) = \frac{\sin(\omega/2) + \sin(\omega/2 - \theta)}{\sqrt{2(\sin^2(\omega/2 - \theta) + \sin^2(\omega/2))}}$$ *Eq. 22 Sine-Law Panning functions*

$$g_R(\theta) = \frac{\sin(\omega/2) - \sin(\omega/2 - \theta)}{\sqrt{2(\sin^2(\omega/2 - \theta) + \sin^2(\omega/2))}}$$

which also satisfy the three constraints Eq. 16 - Eq. 18. The *sine law* is based on the assumption that the listener's orientation is fixed in the direction of the midpoint of the speakers. Pulkki (1997) suggests in place of the *sine law* using the *tangent law,*[138] which has been shown to be more correct when the listener is facing the direction of the virtual sound itself:

$$\frac{\tan(\omega/2 - \theta)}{\tan(\omega/2)} = \frac{g_L(\theta) - g_R(\theta)}{g_L(\theta) + g_R(\theta)}$$

Eq. 23 Stereophonic Tan Law

Similarly, the *tangent Law* yields:

$$g_L(\theta) = \frac{\tan(\omega/2) + \tan(\omega/2 - \theta)}{\sqrt{2(\tan^2(\omega/2 - \theta) + \tan^2(\omega/2))}}$$

$$g_R(\theta) = \frac{\tan(\omega/2) - \tan(\omega/2 - \theta)}{\sqrt{2(\tan^2(\omega/2 - \theta) + \tan^2(\omega/2))}}$$

Eq. 24 Tan-Law Panning functions

Comparing the panning functions, we see that for 12 speakers distributed around 360° and an inter-speaker angle is just 30° ($\omega = \pi/6$), the difference between the s*ine* and *tangent* laws is negligible.[139] In fact, the small-angle approximations $\sin(x) = \tan(x) = x$ for small x, show that as the number of speakers grows, both the *sine* and *tangent* laws converge towards:

$$g_L(\theta) = \frac{\omega - \theta}{\sqrt{\theta^2 + (\omega - \theta)^2}}$$

$$g_R(\theta) = \frac{\theta}{\sqrt{\theta^2 + (\omega - \theta)^2}}$$

Eq. 25 Panning function limit

Note that panning functions in Eq. 25 are simply the linear functions $g_R(\theta) = x$ and $g_L(\theta) = 1 - x$, with $x = \theta/\omega$, and then normalised by the factor described in Eq. 18.

---

[138] (Bernfeld, 1973; Bennet et al., 1985). See also (De Sena et al., 2013) for an alternative derivation of the *tangent law* based on physical aspects of the reproduced sound field.

[139] With 60° between speakers, which is the separation used by Pulkki when proposing the *tangent law*, the difference is more evident, but not significant.

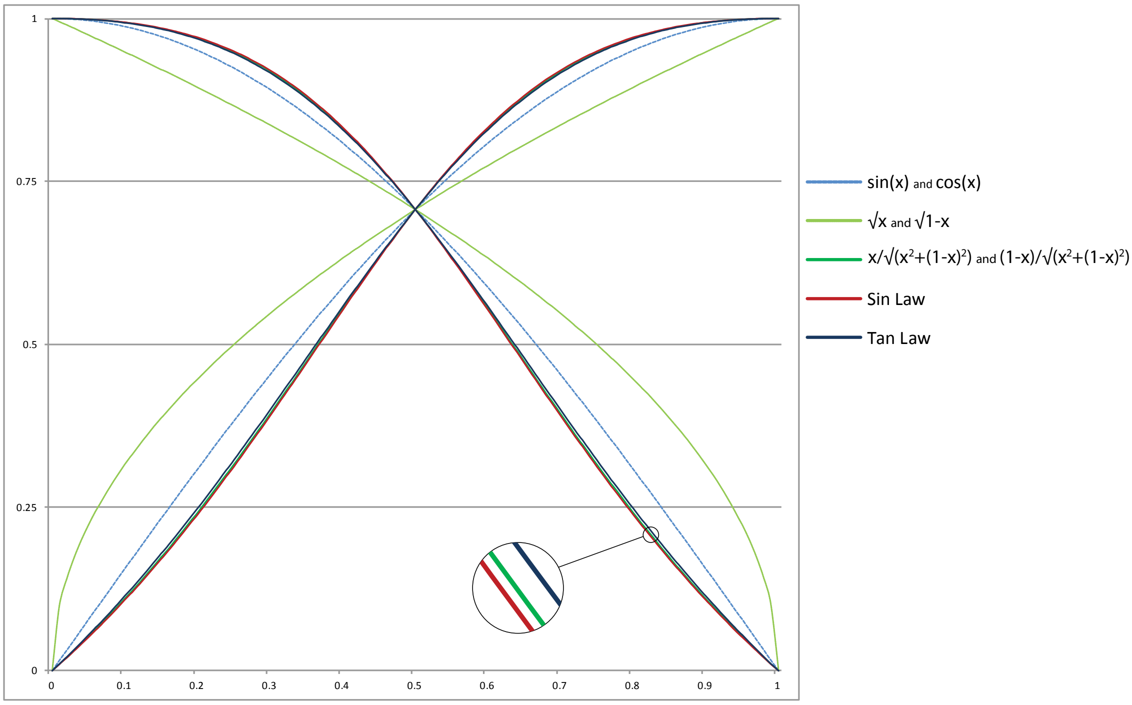


Figure 49 - Panning Functions compared for an inter-speaker angle of 30° ($\omega = \pi/6$)

3.10.4 Uniform Directional Size

With amplitude-panning as it is formulated here, sound sources are modelled as points. However, the perceived angular size of the sound - its 'directional spread' - is not infinitesimal but slightly distributed in space. One reason for this is that perceived direction of a phantom sound varies with frequency (Pulkki, 1999; Griesinger, 2002), and this variation is not captured in any of the panning functions proposed above. A phantom sound composed of a broad spectrum of frequencies will therefore appear slightly spread along the arc between speakers. In addition, the spread of a sound is not constant. When the direction of a sound lies concordant with the position of a speaker, the sound is voiced through one speaker alone and its directional spread is minimal. As the sound pans towards the next speaker, the size of its directional spread increases, reaching a maximum at the midpoint between speakers, before decreasing again as it approaches the next speaker. Pulkki (1999) estimates that with 30° inter-speaker angle, the directional spread ranges from zero at the speakers themselves up to roughly 3° at the midpoint between speakers.

As such, the apparent size of a sound varies slightly as the sound travels between speakers. Pulkki, who would prefer that the sound spread remain constant, seeks to remedy this by introducing *multiple-direction amplitude panning*. Rather than represent a sound as a single point-source, two point sources are used, each displaced an equal amount to the left and right of the desired location for the sound.[140] When the two points fall within the same speaker-pair, the result is identical to that of normal amplitude panning, as the gains sum to the same values that would result from a single sound at the midpoint of the two (see Eq. 26). However, when the points fall on either side of a speaker, the sound is now voiced through three speakers. The result is that a sound travels around the circle, it is voiced through either two or three speakers and never just one, and so always subject to summing localisation. And because the perceived sound is always phantom and never real, it exhibits a more constant directional spread.

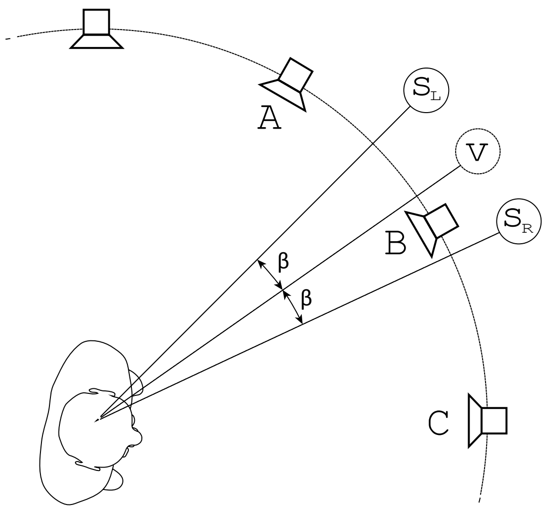


Figure 50 - Multiiple-Direction Amplitude Panning.
Two point S_L and S_R are used to represent single virtual sound v. They are offset to the left and right by the spread angle β, and as a result, v is voiced through 3 speakers A, B and C.

Multiple-direction amplitude panning adds a further constraint on the panning functions as, ideally, the position of the phantom sound should not change with the spread angle. For this to be true, the ratio of left and right gain factors should not alter with spread angle:

[140] Or three point-sources in the three-dimensional case.

$$\frac{g_L(\theta+\beta)+g_L(\theta-\beta)}{g_R(\theta+\beta)+g_R(\theta-\beta)}=\frac{g_L(\theta)}{g_R(\theta)}$$

where θ is the desired sound direction,
β is the spread-angle
$\theta - \beta => 0$ and $\theta + \beta <= \omega$

Eq. 26 MDAP constancy

Of the panning functions presented above, Eq. 19 and Eq. 25 satisfy this constraint. However, neither Eq. 20, Eq. 22, nor Eq. 24 satisfy this constraint, so care must be taken to accommodate changes in direction resulting from changes of spread angle.

With our 30° inter-speaker angle, and the two virtual sound sources displaced to the left and right of the desired sound direction by a 'spread-angle' of 10° each, the peak 3° directional spread observed at the midpoint is now uniformly observed everywhere (Pulkki, 1999). In practice, the 3° oscillations in sound spread are largely rendered unnoticeable in the AVIE by other imaging artefacts, and so multiple-direction amplitude panning is not used.

3.10.5 Software

To facilitate the creation of three dimensional spatial surround soundscapes and compositions, the *icSoundEngine* module was created as part of the iCinema SDK suite.[141] The underlying premise of the sound system is that all sounds and sound events are treated as objects in 3D space, which the programmer or designer controls and manipulates in an identical manner to all other virtual objects in the simulation engine. Behind the scenes, based on the relative positions of the sound objects and the listener (who is also treated as 3D entity), the engine calculates the appropriate speaker levels for each sound using the pair-wise amplitude-panning methods described above. Henceforth, the developer need only concern themselves with the position and movements of the sound objects themselves, and spatialisation is performed automatically.

[141] For a short period prior to the development of icSoundEngine, AVIE applications used a spatial renderer developed by Tim Kreger, based on the *Supercollider* sound platform (http://supercollider.sourceforge.net). *Open Sound Control* (OSC) provided a communication protocol between the Virtools composition on the master computer and SuperCollider running on a dedicated sound computer. This was abandoned for an integrated approach, which benefits from a tighter coupling between the sound and graphics systems, eases debugging and eliminates the need for inter-computer or inter-process communication.

As all sounds are modelled as 3D entities, they have a place in the simulation's hierarchy of 3D reference frames and can therefore be attached to their graphical counterparts, or vice versa. This way, sounds automatically follow their graphical counterparts and are guaranteed to emanate from the same location as their visual representation. Further, events in the virtual world, such as collisions between objects, can be configured to automatically trigger sound effects, and sound parameters such as gain, pitch or reverb can be determined by an underlying simulation. For example, a bouncing ball can be configured to automatically emit a sound at the exact moment and location of impact with a surface, with volume and pitch determined by the force of impact, and other qualities of the sound determined according to the material properties of the surface. Such a simulation-based approach to sound composition greatly simplifies the creation of coherent, synchronous, conflict-free visual and aural cues, allowing the developer to create virtual worlds saturated with the multi-modal spatial cues previously identified as important to presence.

This spatial model also opens avenues to completely different methods of composing non-diegetic sound elements. Sounds without visual counterparts, such as music or narration or ambient effects, can also be treated as spatial narratives. Traditional multi-track mixing of sonic elements, where the relative importance of different voices are controlled by adjusting the volume of each voice over time, can be augmented or completely replaced by a spatial mixing, where the relative volume of sounds are determined by their size and position in space, and the position of the listener.

The spatial model, however, offers more than just a new method of effecting a 'mix', for it constitutes a radically different approach to soundscape or musical composition. Now, music or soundscapes can be composed as spatial narratives – spatial simulations rendered not visually but sonically. With the spatial model, sounds can be arranged as vast structures, constellations or landscapes. They can be put in motion or experienced by a moving listener. Sounds can be triggered by unseen spatial events, such as collisions and interpenetrations, occlusions or revolutions. Arbitrary sound parameters can be tied to arbitrary physical properties of an invisible scene - frequency to the size or speed of an object, reverberation to the distance between two points, or volume to the direction of one object from another. This might all be done by hand, carefully constructing a spatial narrative by traditional animation techniques, for example, but the true power of this method of composition lies in providing the spatial scene with its own kind of forces and behaviours, and letting the spatial narrative unfold autonomously. Like this, soundtracks can easily be given elements of unpredictability, and made to never unfold the same way twice, or come to an end. And like this, soundtracks can be made interactive and reactive to other parts of the virtual world. In *La Dispersion du Fils*, the sound system is used in exactly this way to create a rich, never-repeating and never-ending field of sound through which the listener travels.

icSoundEngine is based on a customised version of the FMODex[142] game audio-engine, modified by iCinema colleague Ardrian Hardjono to support the 12-speaker-system of the AVIE. The icSoundEngine is embedded within the iCinema SDK/*Virtools* engine, complete with C++, VSL and *Virtools* "building block" interfaces. The sound-engine has a number of features to make development easier. For instance, there is a graphical debugger for visualising the state and positions of sounds, and there are mechanisms for asynchronously (i.e. multi-threaded) loading sound-files into memory. One feature that has proven very useful is the engine's capacity to automatically adapt to the number of speakers detected at run-time. This greatly simplifies software development, as no modifications are required to move between systems with two, four, 5.1, or 12.2 speakers.

Now, it is important to note that the sound engine is strictly *sample-based.* It takes as input pre-recorded or pre-synthesised sound files and positions them in space. *It is not a sound synthesiser.* The gain and pitch of a sample can be easily altered. FMODex provides a number of real-time manipulations such as reverb, chorus, low and high pass filters, echo and distortion, but the icSoundEngine cannot synthesise new sounds; instead, all emanations are grounded on a pre-existing sound.

142 http://www.fmod.org

Sound files can be stored in RAM, or streamed from storage dynamically. The number of sounds the engine is capable of emitting simultaneously is limited by the speed of the CPU. While the engine can keep track of tens of thousands of sound-objects, it may only be able to give voice to a fraction of these at any one instant. The exact number depends on the overall CPU load of any particular application. This limitation is managed by allowing the developer to specify the maximum number of voiced sounds at any one time – its *polyphonic limit* -, and having the sound-engine automatically select which sound-objects should be rendered according to priority, which is given to sounds that, after distance and other effects are taken into account, are perceived as loudest. In the experience of the author, giving voice to the loudest 64 sounds has proved sufficient for many applications. The engine also supports sound *instancing*, which is to say that a single source file may be shared between multiple sound emitters, each with unique position, volume, time or frequency, without consuming more computer memory than a single sound. Lastly, in addition to spatialised point sounds, the engine supports non-directional ambient sounds, which are voiced in equal part by all of the AVIEs speakers at the same time.

3.10.6 Distance Perception

Amplitude panning produces phantom sounds that appear to lie on the circle circumscribed by the speakers. In order to produce sounds that appear to emanate from further afield, beyond the perimeter of the AVIE, or closer, within the circle of speakers, different techniques, based on an understanding of how humans perceive distance to sound, are required. The study of auditory depth perception is a complex and rich field of research, with much still to be understood. The reader is directed to Coleman (1963) and Zahorik et al. (2005) for a brief summary of the field. Discussed here are those cues that lend themselves readily to simulation in the AVIE.

The strongest and perhaps simplest of distance cues is the decline in intensity with distance. As a sound grows more distant, it grows quieter. Exactly how volume declines with distance is a function of the shape of the emitted wave-front, and the acoustics of the environment. For an idealised point-source emitting a spherical wave-front in an acoustic-free field, sound pressure is proportional to the inverse of distance 1/d.[143] This is the default behaviour in the icSoundEngine. However, there are times when the sound designer may wish to deviate from this ideal. For example, with sound waves emanating from a linear or cylindrically shaped emitter, sound pressure declines with $1/\sqrt{d}$ rather than 1/d. Large-scale acoustic elements, such as the ground or a long corridor may also alter this decay, by directing more of the emitted sound energy in the direction of the listener. These various declines in volume with distance are easily simulated in the icSoundEngine. Sounds are modelled with a minimum radius, within which the volume is held at a constant maximum, and beyond which they decline according to some user-specified distance-decay function.

Second, as sound travels through air, high frequencies are attenuated differently to lower frequencies, producing a change in the sound's frequency spectrum. The nature of the change is complex, depending on such factors as humidity, temperature and pressure, but in general it is the high frequencies that are most attenuated (Coleman, 1963). As such, a suitably tuned low-pass filter can serve as a simple approximation of this depth cue.

When a virtual event is both visible and audible and sufficiently distant, the delay between the arrival of visual and sonic stimuli caused by the different velocities of light and sound can serve as a potential distance cue. As distance increases, sound waves arrive further behind their visual counterpart, in a manner that is easily simulated by delaying the audio one third of second for every 100 metres to the sound source. It should be noted, however, that if any of the sounds are continuous or enduring, rather than short, sharp emissions, and the virtual listener or sound sources are in motion at all, then this distance cue also demands the simulation of Doppler-shift, for this is the only way of effecting a smooth change in time-of-flight delays as distances between the virtual listener and the virtual source change. While FMODex does furnish the icSoundEngine with a real-time Doppler shift filter, the sonic quality of the pitch shift produced is extremely poor and thus largely unusable. As such, disparity between

[143] Sound intensity follows the inverse square law, while sound pressure is proportional to the square of intensity.

time-of-arrival of sound and vision must be used judiciously as a depth cue in dynamic scenes.

Another cue employed by human hearing to judge distance is the *direct-to-reverberant energy ratio*: the ratio of sound energy reaching the listener directly to that arriving via reflections, reverberations, diffractions or other interactions with the environment. The closer the sound source, the greater proportion of energy received directly, while the further the sound, the greater the energy received via reverberation. As such, modulating the ratio of direct sound to reverberated sound is an effective method of inducing the sensation of distance (Bronkhorst & Houtgast, 1999). To employ this method, however, some model of the reverberant qualities of the virtual environment is required. This might be grounded in simulation, with reverberations, reflections and diffractions calculated directly from geometrical representations of the virtual world (e.g. Antani et al., 2012), or very loosely approximated by one or more zones of 'high' or 'low' reverberation. Ideally, the direct component of the sound would possess a very precise sound direction, while the reverberated component would be dispersed, arriving from multiple directions.

Finally, one of the more important distance cues is provided by familiarity with the sounds themselves. For example, when the same utterance is recorded twice, once as a whisper and once as a shout, and each of them replayed with equal intensity, the whisper is consistently perceived as closer than that of the shout (Brungart & Scott, 2001). Familiarity with a sound allows an estimation of its 'true' loudness which when compared with perceived loudness, provides a measure of distance (Philbeck & Mershon, 2002). As Zahorik et al. (2005) puts it:

> Listeners are likely to have some long-term knowledge about the characteristics of most naturally-occurring sound sources, and this *a priori* knowledge is likely to influence how they judge the distances of these everyday sounds.

This leads to the final point, which is that distance cues can be, and often are unavoidably, encoded within the samples themselves. Field recordings, for example, are typically rich with complex distance cues such as reverberations and spectral attenuations that accurately convey the distance from the sound to the microphone as well as the acoustic character of the environment. Such sounds are often described as 'wet', which is to say that environmental cues (reverb and spectral attenuation) are encoded within the sound samples. In contrast, a 'dry' sound is one in which these cues are absent, the sound having been captured with a close microphone in a controlled, anechoic room, or having been artificially synthesized. Dry sounds are necessary starting points if the sound is to undergo real-time reverberation or spectral filtering, for once these cues are encoded within a sample they cannot be removed.

There are, therefore, two different approaches to the simulation of distance with the icSoundEngine. The first is to use dry sounds, and apply one or more of the effects described above in real-time. Alternatively, a number of 'wet' versions of the same sound, identical save for the apparent distance to the listener, might be prepared outside of the real-time simulation. These different versions of the sound might be captured with multiple microphones placed at different distances from a real sound source. Or, because the computations are only performed once and need not be real-time, different versions of the sound might be pre-computed with slower, more computationally expensive but more accurate algorithms than those possible in a real-time simulation. Both methods are viable. Then, during the real-time simulation, the sound engine simply mixes the different samples according to the distance of the virtual sound source.

The advantage of the 'wet' approach is that it allows for a far higher quality of reverberation and spectral attenuation than might be possible with a real-time simulation, and dramatically reduces the computational workload of the real-time simulation, thereby allowing for the sonification of more sounds with the same computational resources. It does so, however, at the expense of flexibility, and if listener is moving through many different acoustic environments, or the environment itself is undergoing significant change, it also sacrifices accuracy. However, with the current state of real-time acoustic manipulations, the pre-computed or pre-captured 'wet' approach tends to produce more realistic or compelling results than the real-time 'dry' approach.

3.10.7 Sound Size and De-correlation

Within the icSoundEngine, sounds are modelled as spheres. However, although the size of this virtual sphere is used to modulate the intensity and decline of intensity with distance, it is not reflected in the perceived size ('directional spread') of the sound source, for with amplitude panning all sounds are perceived as emanating from a point.

Very often, this is perfectly appropriate; the words emanating from a person's mouth, the sound of footsteps or clapping hands, a door slamming, a gun firing, or the sounding of bells are all examples of phenomena that can be well represented by a point source. At other times, however, this is inappropriate. A passing train, a forest stirred by wind, the chanting of a vast crowd or the chatter of a large flock of birds, when heard from nearby or within, are all poorly represented by a single punctual sound.

It is not difficult to see that these phenomena are all collections of sound-emitting entities - trees, people or birds - and therefore ideally modelled as multiple point sources; one for each tree, person or bird. The train too is best modelled as a collection of parts, with a point source for each wheel, bell, whistle or clanging chain.

Breaking down large sonic objects into structures of point sources is a highly effective means of sound composition, and it is used to great effect in *La Dispersion Du Fils*. Very often it can be achieved by simply 'attaching' sounds to the appropriate sub-components in the same 3D hierarchy of the scene employed by the rendering system. However, there are a number of subtleties requiring attention.

The first is that of coherency. Coherent sounds, regardless of their number and position will, due to summing localisation and the precedence effect (discussed below), be perceived as a single phantom sound. *The very same perceptive mechanism exploited in amplitude panning to provide sounds with direction is now working against us*. To prevent this perceptual fusing, the individual sounds must be sufficiently different from one another that they are not subject to summing localisation. Each bird must be given a unique song, each person in the crowd a unique voice, each leaf its own rustle, and each wheel on the train its own particular clickety-clack.

In practice, sourcing a unique sample for each sonic element - each bird, rustling leaf, member of the crowd or wheel of a train – can be impractical or unfeasible. In some cases this can be avoided by using a single long sample for all sound atoms, but giving each individual sound atom a unique temporal offset or shift in frequency. The viability of this technique depends greatly on the type of phenomena being represented; on whether it has temporal features that betray repetition or whether the sounds must maintain synchronisation with one another or some visual timeline. For example, one long recording of a rustling tree, voiced multiple times at different positions with significant temporal shift and slightly different pitch at each position, may well suffice for the simulation of a forest. For the singing crowd, however, it is clear that significant shifts in time or frequency are desirable.

Another approach is to use decorrelation algorithms to produce multiple decorrelated copies of a sound. Decorrelated sounds are soundwaves that are indistinguishable on a macro-scale, and so sound identical, but differ significantly on a micro-scale (of the order of wavelength), and so are resistant to summing localisation, the precedence effect, as well as constructive or destructive interference (Kendall, 1995; Liu & Smith III, 2002; Potard & Burnett, 2004). A typical approach is to pre-compute a library of decorrelated copies, or potentially a real-time decorrelation technique might be adopted (e.g. Cecchi et al., 2012). The two approaches draw on the computing system's resources differently, with the former taking more memory and disk activity and the latter consuming significantly more CPU cycles.

A second challenge is that, while a typical computer may be able to keep track of thousands of point sounds, it will only be capable of giving voice to a small number of sounds at any one time. Suppose that only a small amount of CPU time can be devoted to sound processing and the polyphonic limit is fixed at just 32. The sound of 32 birds, leaves or people is clearly perceptibly different from the sound of a thousand birds, leaves or people.

A possible solution is pre-compute mixes of individual elements. Supposing, for example our flock is composed of 1000 birds, then rather than use one sample for each bird, 10 samples each containing the song of 100 birds might be used. The challenge that follows is to position and move these 10 sounds in a manner that is representative of the movement of the birds as a whole. A naive approach would be to attach the sounds to every 100th bird and rely on the statistical spatial uniformity of the flock. Alternatively, one could continually group the birds into clusters (for example, using the *k-means* algorithm or similar (Hartigan, 1975)), and position a sound at the centre of each cluster. A more subtle approach would be to pre-compute mixes with varying numbers of voices (for example 1, 2, 4, 8, 16... 512) and then cluster the birds according to distance and angular size from the perspective of the listener, and distribute the appropriate pre-computed mixes to each cluster. Nearby birds would be given each their own voice, while birds slightly further afield would be represented as groups of 3 or 4, and so on. This ensures that any birds in close proximity to the listener are never voiceless, an aberration that would be highly disruptive to presence.

3.10.8 Accuracy and Distortion

In formulating the panning functions above, two assumptions were made that do not always hold true. The first is the assumption made in the drafting of the stereophonic tangent law (Eq. 23) – that the listener is facing the phantom source. In the AVIE, with sound emanating from all directions at once, and with multiple listeners turning freely and independently, this assumption is almost always false. The second is that the listener is fixed in the centre of the circle. With multiple listeners free to move about within the AVIE, the assumption of a centrally fixed listener is also false.

The Precedence Effect

To understand the ramifications of these false assumptions, it is necessary to first understand in greater detail the phenomenon of summing localization and the mechanisms by which sound localisation is achieved in human hearing. Sound localisation is afforded by four different cues (Pulkki & Karjalainen, 2008):

- The interaural time difference (ITD); a measure of the phase and time difference between signals arriving at each ear, arising from differences in propagation time from the sound source to either ear.
- The interaural level difference (ILD); a measure of the difference in intensity between signals at each ear, arising principally from the shadow cast by the head and ears.
- Changes in the spectrum of the sound due to filtering effects of the ear, head and body.
- The effect of head motion on all of the previous three cues.

The first two cues can be considered the primary cues, with the ITD taking precedence below 1.6 kHz and the ILD above 1.6 kHz. However, the position of a sound cannot be completely derived from the ITD or ILD alone, for with any given ITD or ILD there exists a "cone of confusion," a cone of sounds all sharing the same ITD and ILD. To locate a sound within the cone of confusion, the third and fourth cues are called upon. Note that the fourth cue is another example of the importance of active perception and the role of ego-motion in perception.

So our ability to perceive direction in sound is primarily due to sound travelling different paths to each ear and arriving at different times. The soundwaves arriving at the left and right ear are perceptually fused into a singular *directional* sound, with the ITD and ILD determining the perceived direction. Amplitude panning is able to exploit this feature of human hearing because the waves arriving from all loud speakers *at each ear* are superposed into a single wave, and the phase of this fused wave is determined by the relative intensities of the wavefronts. This is the counter-intuitive mechanism underlying amplitude panning: the level difference between the loudspeakers affects the interaural time difference (ITD).

We name the first sound to arrive at the listener the *lead*, and the second to arrive the *lag*. For summing localisation to take place, the time between reception of lead and the lag stimuli *must be less than 1/1000th of a second*. Only when the time difference is below this threshold, summing localisation may occur (Warncke, 1941; Blauert, 1997; Litovsky et al., 1999).

As the time delay approaches 1 millisecond,[144] the apparent position of the phantom sound shifts gradually towards the position of the lead speaker. At 1 ms, an effect known variously as the *precedence effect* (Wallach et al., 1949) or *law of the first wavefront* (Cremer, 1948) or *localization suppression* (Litovsky & Shinn-Cunningham, 2001) takes hold. Here, the two sounds continue to be perceived as a single acoustic event, but the *apparent position of this event is now entirely dictated by the position of the lead sound*. The lag sound continues to effect the perception of the sound, but only by altering the perceived spatial size and "tone colour" of the event. It no longer plays any role in the perceived direction of the sound.

The purpose of the precedence effect is to distinguish between delays introduced by the different paths around the head, and delays introduced by reverberations and reflections within the environment. With a separation of 21cm between the ears and a speed of sound of 343 m/s, sound waves following direct paths from a true point source can arrive at one ear no later than 0.6ms than the other. A delay greater than 0.6ms between the lead and lag is never caused by the longer path around the head, but can only arise from sound waves travelling an indirect route to the ear. In order to prevent such indirect routes (reflections, dispersions, refractions) influencing the perceived direction of the primary sound, they are excluded from the calculation of direction. This is the precedence effect.

The precedence effect is the mechanism that allows us, in the real world, to perceive a singular acoustic event as taking place in a singular place and time, despite arriving at our ears as a complex constellation of reflections and reverberations. The effect is therefore important to our ability to understand and navigate acoustically chaotic and complex sound environments.

[144] 1 ms is the most frequent value cited in literature, but the shift is non-linear. Harima et al. (2013) provide a recent study confirming Blauert's (1997) observation that the movement of the phantom sound towards the lead position is almost complete after just 0.63 ms. While this varies a little with changes in the relative gain levels of the two sources, it is clear that this value reflects the maximum possible ITD arising from the geometry of the head alone. For more results please see Harima et al. (2013)

As the time delay between the lead and lag sound rises above 1ms, it approaches a new threshold called the *echo threshold* (Litovsky et al., 1999). Here, fusion begins to break down, and the lead and lag sounds are perceived as separate sound events. The echo threshold varies enormously with the character of the sound, ranging from 5ms - 10ms for short clicks, to 30 - 40ms or longer for speech or music.[145] It also varies with the difference in intensity between the lead and lag. Note, however, that while the lag sound is now perceived to clearly occur after the lead sound, its apparent position continues to be determined by the lead sound. This phenomenon is known as *lag discrimination suppression* (Litovsky et al., 1999; Litovsky & Shinn-Cunningham, 2001). As time delays further increase, lag discrimination suppression eventually subsides, and the lag sound is finally perceived as an echo, with its own clearly discernible positions in space.

Turning Listener

Whenever the listener turns, the relative arrival times of all four wave-fronts are altered, and the relationship between relative gain and perceived direction described in the tangent and sine laws no longer holds true. Theile and Plenge document this phenomena with a circle of six speakers (1977), and their results are reproduced in Figure 51 and Figure 52 for illustrative purposes. Figure 51 below shows how the perceived position of the phantom sound is accurately described by the tangent law when the active speaker pair is directly ahead.

[145] See Litovsky et al. (1999) for a review of echo thresholds in literature. The most commonly cited lower limit for the echo threshold for short clicks is 5ms-10ms with the "lag clearly audible on 75% of trials."

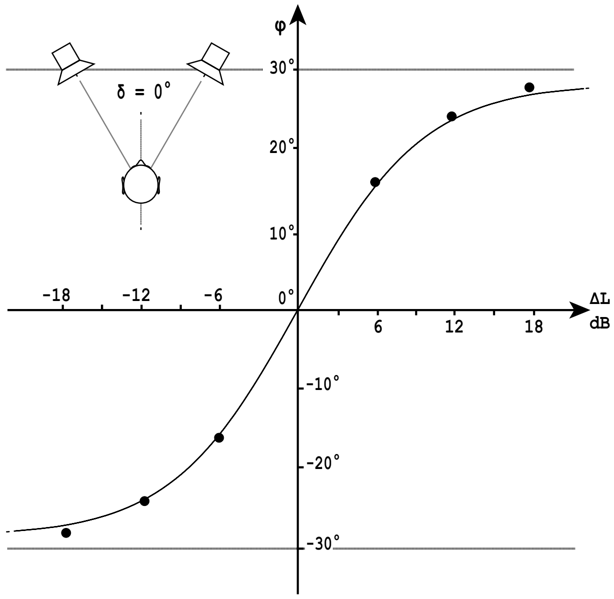


Figure 51 - Perceived Position of sound for front facing listener δ = 0°.
Note how the Tangent Law (depicted by the smooth curve) accurately describes the real-world measurements (the points on the curve). (Reproduced from Theile and Plenge (1977)).

In Figure 52a to Figure 52d, the effects of listener rotation are depicted. As the listener turns and the active speaker pair shifts from the front to the side of the listener, the phantom sound is pulled towards the speaker that is closest to the front or back of the listener (Theile & Plenge, 1977; Pulkki, 2002). This is the precedence effect asserting its influence. As the slope of the curve grows, the perceived position of the sound becomes increasingly sensitive to movements of the listener. A virtual sound moving with constant velocity will no longer be perceive as such, but tend to slow down near speakers and travel fastest at the midpoint between speakers. Also, the curves no longer pass through the origin, implying that equal speaker volume no longer produces a phantom sound at the midpoint of the two speakers. Note also the increasing error ranges, which reflect a decrease in stability of the phantom sound.

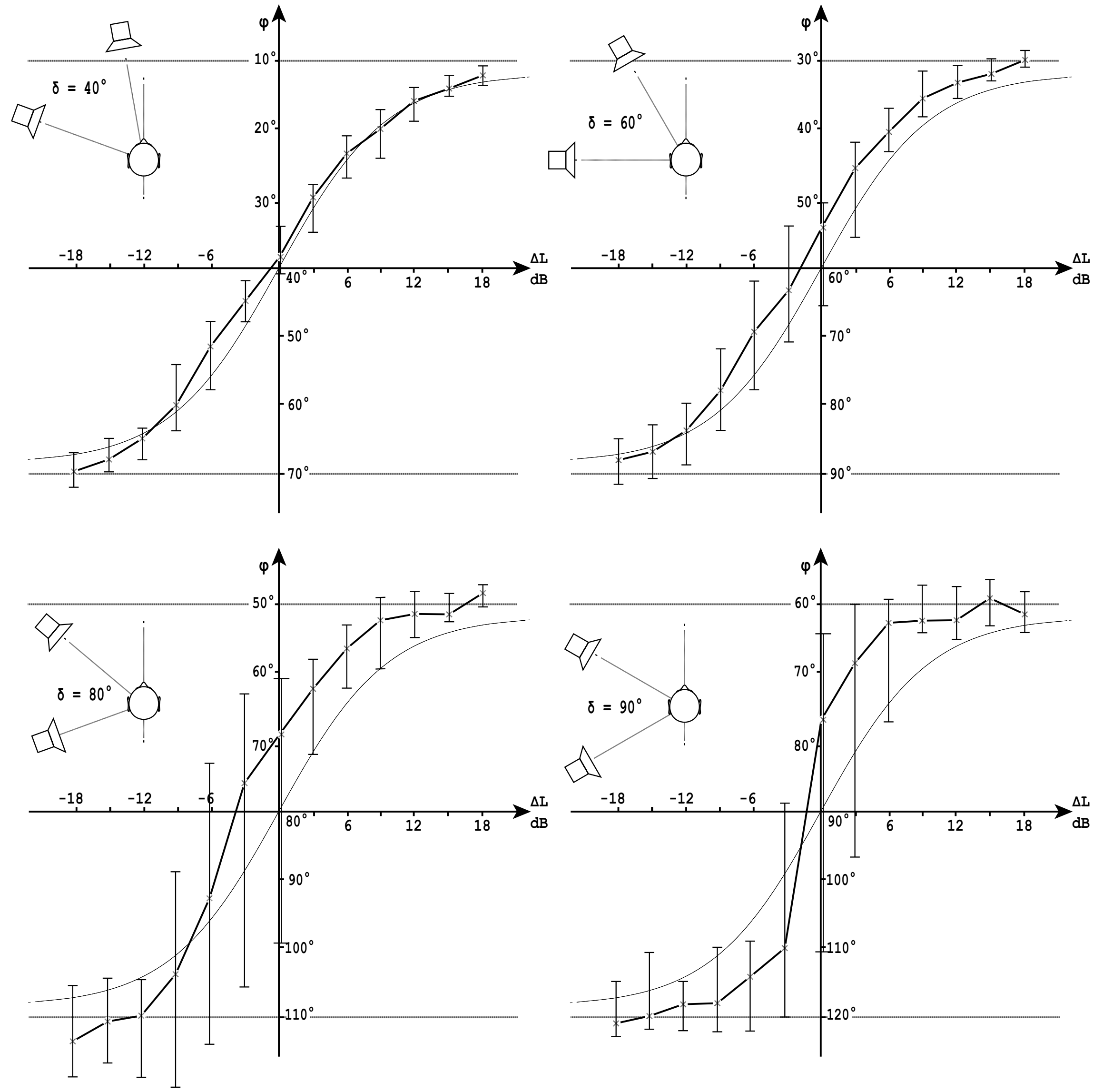


Figure 52 - Perceived direction for rotated listener a) $\delta = 40°$ *b)* $\delta = 60°$ *c)* $\delta = 80°$ *d)*$\delta = 90°$.
(Reproduced from Theile and Plenge (1977)).

When the two speakers are directly to the side of the listener and symmetrical around the inter-aural axis (Figure 52d), they both share the same "cone of confusion." Here, the phantom sound position is at its most unstable, shifting between one speaker and the other with the subtlest of changes in head orientation, or the smallest of changes in gain ratio. Note how the error range in the centre of Figure 52d, where the virtual sound is desired to be exactly in the middle, spans from 65° to 110°, three-quarters of the inter-speaker distance.

Amplitude panning, therefore, is most accurate and stable when the listener is facing a pair of speakers, and at its most unstable and inaccurate towards the side of a listener. It should be noted that part of this degradation in localisation is due to the natural decline in our capacity to judge the direction of the sound in the periphery. As such, the lack of peripheral spatial accuracy inherent in amplitude panning is somewhat masked by our natural lack of spatial accuracy in these directions (Pulkki & Karjalainen, 2008).

The Sweet-spot

One of the key properties of a spatialisation system is the sensitivity to the position or orientation of the listener. In audiophile parlance, the "sweet-spot" describes the zone within which the accuracy of localisation is acceptable. For amplitude-panning array, the sweet-spot is the zone within which summing localisation takes place, which was identified above as the zone within the AVIE within which the temporal difference between lead and lag does not exceed 1ms. In the AVIE, with 12 speakers distributed around the 31.4m circumference, each speaker is 2.6m apart. Therefore, with a speed of sound of 343.2 m/s the time delay between lead and lag from neighbouring speakers can be, at its maximum, 7.6ms.[146]

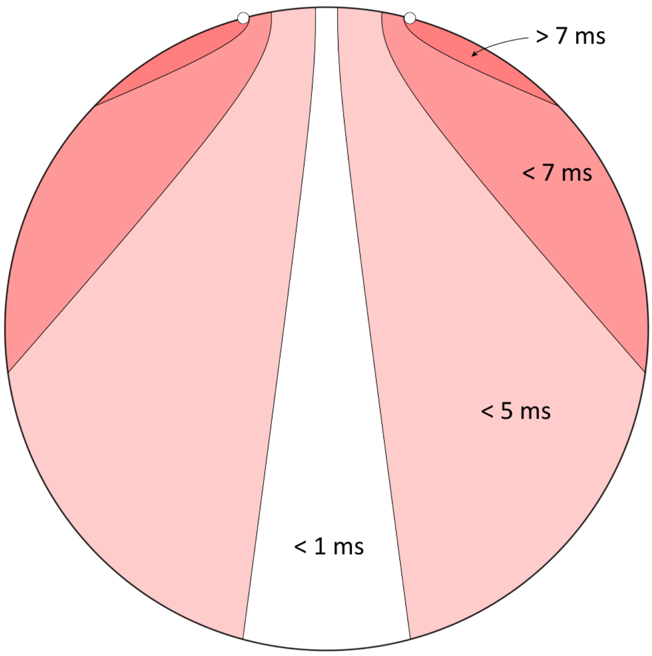


Figure 53 - Lead-lag delay for a single speaker pair.

[146] 343.2 m/s in dry air at 20 °C

Figure 53 depicts the difference in time of flight for the lead and lag emanating from neighbouring speakers for all points within the AVIE. The narrow corridor through the centre is the region within which the delay between lead and lag is less than 1ms, within which summing localisation takes place. Outside of this 'summing corridor,' the precedence effect dominates, and a sound is perceived to emanate from the closest speaker of the pair. The second contour marks the 5ms boundary, which has been cited as the lower limit of the echo threshold for short click stimuli. At 5ms, it is possible that short clicks are no longer heard as single sounds. However, this is the lower limit and only for a very particular form of sonic event. The 7ms threshold is shown also, and effectively marks the maximum possible delay between lead and lag in the AVIE.

Figure 54 shows the maximum lead-lag delay at all points in the AVIE, for *all* pairs of neighbouring speakers. The inner-most circle, in which the delay is less than 1ms for all directions, can be considered the sweet-spot of the AVIE.

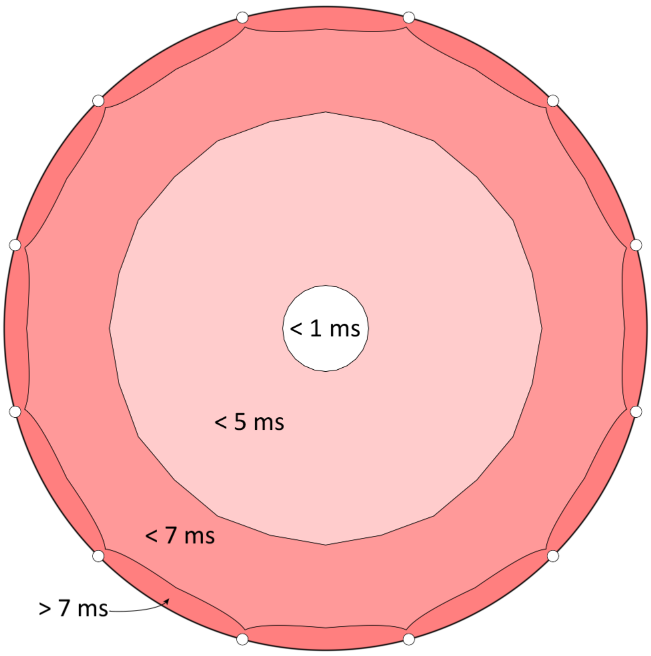


Figure 54 - Maximum lead-lag delay for all neighbouring pairs of speakers.

As can be seen, the sweet-spot for the AVIE, within which amplitude panning behaves as expected and the perceived position of the phantom sound matches its virtual source for all speakers pairs, is a small circle of just 1.3 metre diameter in the centre of the AVIE. (And this is somewhat generous; a figure of 0.6ms / 0.8m would be more accurate). When the listener ventures out of this zone, and depending on the orientation of the listener relative to the sound, the precedence effect comes into play and phantom sounds are drawn towards the closest speaker to the listener of the pair.

The size of the sweet-spot is directly proportional to the number of speakers. At 24, the sweet-spot doubles to 2.7 metres, and at 48, it has reached 5.2 m. With 96 speakers, at no point within the AVIE is the difference between the lead and lag greater than 1ms. Note that at 96 speakers, wave-field synthesis or higher-order ambisonics can begin to be considered as viable alternatives to amplitude panning.

Distortion

As the listener moves away from the centre of the AVIE, the perceived location of phantom sounds grows more discordant with the position of the sounds in the virtual world. This distortion has a number of causes. The first is that introduced by the assumption of a single central listener when calculating the azimuth of the virtual sound position. The directional error ϕ_ε introduced by this assumption can be written as a function of the position of the listener (radius r and direction θ), the radius of the AVIE and the distance of the virtual sound:

$$\phi_{\varepsilon} = tan^{-1}\left(\frac{d - r\,sin\,\theta}{r\,cos\,\theta}\right) - tan^{-1}\left(\frac{R - r\,sin\,\theta}{r\,cos\,\theta}\right)$$ *Eq. 27 Geometric Azimuth Error*

where R is radius of AVIE,
r and θ are the position of the listener with the AVIE,
and d the distance to the virtual sound.

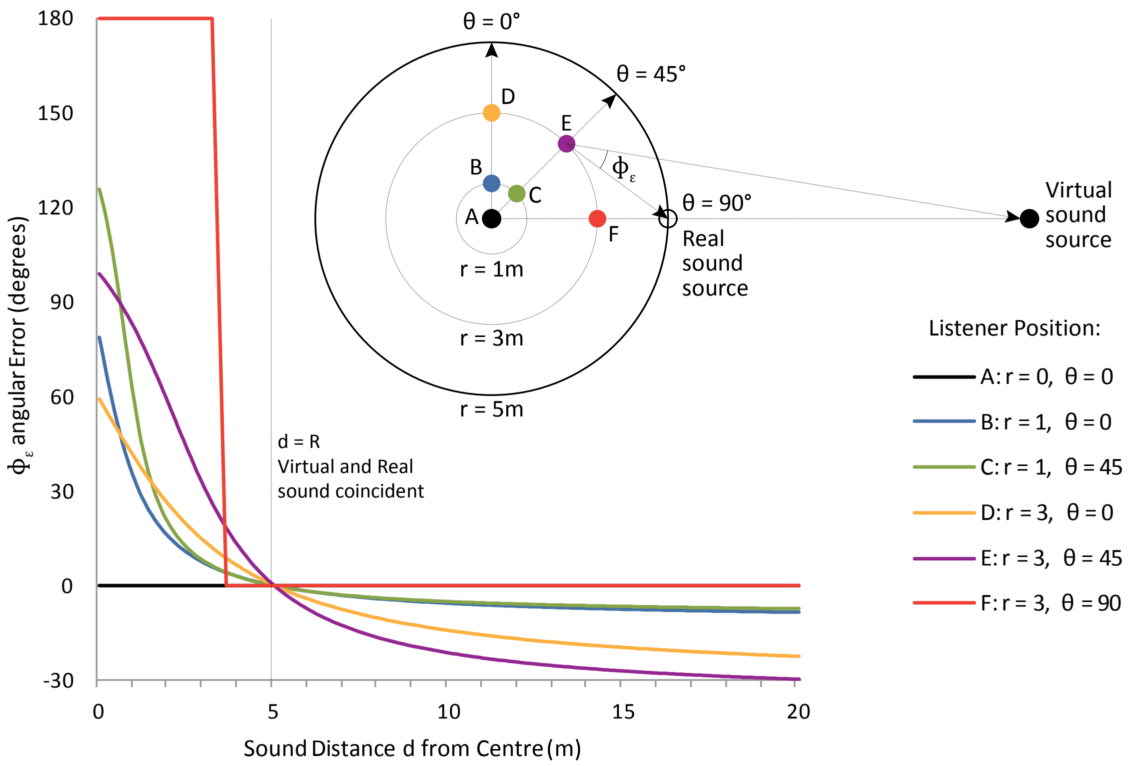


Figure 55 - Directional error introduced by the central listener assumption.

In Figure 55 the directional error ϕ_{ε}, as defined in Eq. 27, is plotted for various listening positions within the AVIE. Note how, for virtual sounds outside the AVIE, the directional error quickly approaches a limit. As the virtual sound distance grows large (d >> R), this limit can be written as:

$$\phi_{\varepsilon\, max} = \frac{\pi}{2} - tan^{-1}\left(\frac{R - r\sin\theta}{r\cos\theta}\right)$$

Eq. 28 Maximum Azimuth Error for d >> R

For example, for a listener within 3 metres of the centre listening to virtual sounds emanating from outside the AVIE, the error ϕ_{ε} is bounded by 36°.

In contrast, for virtual sounds within the AVIE, this error rapidly grows and without bounds. For this reason, the system is ill-suited to the simulation of sonic events within the perimeter of the AVIE.

In addition to this distortion is the persistent shift of the phantoms sounds towards the closet speaker, described in detail above. This can be considered a form of high level *spatial aliasing*, where sounds tend to converge at the physical speaker locations.

When the listener moves laterally across the summing corridor, a phantom sound moves from one speaker towards the other. The stability of the phantom sound is therefore proportional to the width of the summing corridor at the point the listener traverses it. If we compare the distance moved by the listener with the resulting shift in direction of the phantom source, it is evident that the stability of the phantom sound increases as the listener moves away from a speaker pair and decreases as they approach it.

Distance perception is also perturbed by a wandering listener. Here, the inverse square-law between intensity and distance which we exploited to create a sense of distance, works against us. For a listener moving within the AVIE, sounds grow and diminish at rates determined by the physical distance to the loud speakers, which, except for the case where the virtual sound lies exactly on the AVIE perimeter, will be incommensurate with their virtual distance. All sounds, regardless of their virtual distance, grow and diminish as if they were only 5 metres away.

3.10.9 Graceful Decay and Effect of Visual Stimuli

Judged in purely numerical terms, the sound system presented here is not without significant flaws and limitations. The sweet-spot is so small that most viewers will never find themselves within it, outside of this region directional accuracy degrades, eventually arising at a form of spatial aliasing where sounds collapse to speaker locations. The system is limited to the simulation of point sounds only. It is unable to simulate sounds arriving from above or below, or from within the AVIE perimeter, and can only crudely simulate the illusion of distance beyond.

Despite all this, the system is extremely effective. Over the many years the AVIE has been in operation, no one has ever noticed a discrepancy between the direction, size or distance of a sound and its visual counterpart. The perceived experience seems to contradict the numerical analysis presented above. So much so, that it beckons an explication of why none of the flaws or limitations forecast by the analysis above are perceptible in practice. There are a number of reasons this is so.

Firstly, the assumption of the central spectator yields the same geometric distortion for both panoramic images and sound, so while sounds may not be perceived as emanating from their true virtual position, neither are their visual counterparts and, most importantly, they share the same distorted position.

Secondly, nowhere in the AVIE is the delay between the lead and lag sufficient to transgress the echo threshold. This means that nowhere in the AVIE does binding of two sounds fail completely where they are perceived as two sounds coming from two different speakers.

Thirdly, the distortions are always smooth and continuous. Although the precedence effect tends to pull sounds towards the closest speaker, which as the listener moves about may change in an instant, the shift of the phantoms sound from one speaker to another is never abrupt but always smooth and gradual. Just as sharp discontinuities in visual distortions diminish presence, jumps or breaks in the sound field must also be avoided if presence is to be preserved.

> We “filter” our perception based on an assumption of continuity. Anything conflicting this assumption will lead a sense of disbelief or a lower sense of immersion or telepresence. (Dickins, 2003, p. 67)

Further, and most importantly, the directional error induced by the precedence effect is bounded in magnitude by the speaker pair:

> In any listening position, the virtual source cannot be perceived outside the sector defined by the loudspeaker pair and the listener. (Pulkki & Karjalainen, 2008, p. 121)

The angle between two speakers provides an upper bound on the directional error. For the centrally located listener, this is 30°.

It is well documented that vision can influence the perception of auditory space, including the perceived direction and distance of a sound. The “ventriloquism effect,” or “visual capture” is the name given to the irresistible shift in a sound’s perceived direction towards a plausible visual origin (M. B. Gardner, 1968; Mershon et al., 1980; Alais & Burr, 2004; Charbonneau et al., 2013). Jack and Thurlow (1973) document the ventriloquism effect producing angular shifts in direction over 30°. This implies that, for large parts of the AVIE where the angle between speakers is not greater than 30°, the ventriloquism effect can completely counteract the distorting influence of the precedence effect, and all sonic emanations with clearly visible origins in the virtual world will be perceived correctly as emanating from their visual counterpart. Similar effects are also documented for the perception of distance (Zahorik, 2001, 2003).

The interplay between vision and hearing, or "cross-modal enhancements" as they were referred to previously, helps explain why the spatial infidelities of the sonic reproduction are essentially, in practice, imperceptible. The limited sweet spot and directional accuracy of amplitude panning have presented no impediments to the effective spatialisation of sound within the AVIE, and the construction of immersive experiences. It is postulated here that the ventriloquism effect and subsequent successful binding of visual and aural cues play a significant role in this. For more on the interplay between vision and hearing in virtual environments, see Storms (2002).

3.11 Computers, Clustering and Synchronisation

The technique of combining multiple projectors or screens into a single unified high-resolution or high field-of-view display poses a very elemental problem: what is to be done when there are more displays than can be connected to, and driven by, a single computer?[147] The solution adopted here is to use multiple computers in a tightly synchronised cluster.

Background

The use of a cluster of PC computers to drive multiple display devices in unison has its origins in the late 1990s. Prior to this, real-time control of multiple displays was strictly the domain of high-end graphics "super-computers" such as the Silicon

[147] The number of displays that can be physically connected to a single computer changes with every generation of graphic cards. In 2005, at the time development of the AVIE began, a single computer using 3 graphics card might support up to six displays, albeit with some performance limitations. In 2014, a single workstation equipped with 4 graphics cards can support 16 or even 24 displays.

Graphics *InfiniteReality* system (Montrym et al., 1997). The prohibitive cost of such machines at least partly spurred the search for an alternative, and it was with the arrival of the commodity PC 3D graphics processing unit (GPU), originally intended for the computer game player, that such an alternative become available. By running a number of PCs, each equipped with a GPU, in a tightly synchronised cluster, multiple displays could be driven at a fraction of the expense of a monolithic super-computer. By 2002 a number of different visualisation cluster systems had been demonstrated (Tramberend, 1999; Allen Bierbaum et al., 2001; Bues et al., 2001; Chen et al., 2001; Stoll et al., 2001; Allard et al., 2002; Humphreys et al., 2002; Voß et al., 2002; Allard et al., 2003; Schaeffer & Goudeseune, 2003; Streit et al., 2004), and soon afterwards clusters of PCs would come to be considered the norm for visualisation systems. For a general overview of the various issues involved in graphics clusters, the reader is directed to L. Soares et al. (2008).

3.11.1 Cluster Design

The AVIE cluster consists of one master computer and six render-nodes. Each computer is a regular PC[148] running the Windows operating system.[149] The computers are connected via an ordinary Gigabit ethernet hub. Each render-node is connected to two projectors operating as a stereo pair (in the case of passive stereo) or a single active-stereo projector. Note that the active-stereo versions of the AVIE may have only six projectors, but each projector requires two display inputs,[150] so in both active and passive configurations, every render-node produces two video signals. As such, in terms of system structure, the passive-stereo and active-stereo versions of the AVIE are identical.

[148] Across the 12 AVIE implementations, the specifications of the PC vary according to the date of their construction. Some systems have dual-CPU computers, while others only one. All of them are Intel based Windows PCs.

[149] Originally Windows XP, later upgraded to Windows 7

[150] These two inputs are then buffered and presented sequentially by the projector in order to produce frame-sequential stereo.

An AVIE configured for development is equipped with a video matrix, through which all video signals are routed. Used with four conventional displays and a keyboard/mouse router, this allows a technician or developer to easily switch between and control any two computers at any time. An additional monitor/keyboard/mouse terminal is provided on a trolley to be easily wheeled in and out of the cylinder, or operated just outside the screen entrance.

The master computer is connected to the sound system described in Section 3.10. Input devices also connect directly to the master computer. The tracking sub-cluster (Section 3.12.1) is connected to the same network.

A small control computer with custom software[151] allows remote control of lights, booting and rebooting of computers and remote operation of the projectors, video matrix and keyboard switch. All computers have access to a centralised file-server. In general, however, high-bandwidth data, such as video and audio, is mirrored on local hard-drives to guarantee fast access. For this, custom utilities for keeping files synchronised across the cluster have been developed.[152]

Active-stereo configurations require genlock synchronisation, provided by Quadro G-Sync cards that lock all the graphics cards in the cluster to a single clock. The scan lines of all displays and buffer-swapping are then guaranteed a very high level of synchronisation.

3.11.2 A Faustian Pact

The benefits of clustering do not come freely. The price extracted for the ability to connect an arbitrary number of displays and the augmented computing power required to drive them is that the software generating the images must now somehow be distributed across multiple computers. The ramifications of this are significant, for clustering immediately rules out the use of existing software unless this software has

[151] Developed by iCinema engineers.

[152] Written by iCinema engineer Robin Chow.

been designed *explicitly* to run across a cluster.[153] If we consider the cluster as a single virtual computer, then it is a computer incapable of running almost all existing software. Apart from a handful of commercial software products equipped with cluster-support, a cluster is restricted to running bespoke, custom software.

This is the first price paid. The second price paid is a significant increase in the complexity of developing said bespoke software. The added complexity comes from two related issues - data distribution and synchronisation - the subjects of the next sections.

3.11.3 Parallelisation and Data Distribution

There are a number of different approaches to the organisation and distribution of software and data across the cluster.

With one approach, a single 'master' computer would be responsible for all computation and rendering while the 'slave' render-nodes serve as nothing more than display devices. In the case of the AVIE, the master would generate 60 times every second a complete 360° stereo panoramic image; and then partition and transmit this image to the render-nodes, which simply pass it through to the projectors. The advantage to this approach is the relative simplicity of the software. The render-nodes act as nothing more than 'dumb terminals,' and apart from the transmission of the images each frame to them, the software on the master differs very little in structure from that of a conventional single computer/single display arrangement.

The disadvantages to this approach are two-fold. First, the amount of data distributed each frame can easily be prohibitive. In the case of the AVIE, transmission of the twelve 1400x1050x24bit images at 60Hz would amount to the movement of 3028 MB every second, well in excess of the ~100MB/s offered by Gigabit ethernet. This could be

---

[153] Cluster-support is not the only factor restricting the use of existing software. Software for AVIE must also support distortion correction, blending, rendering of stereo panoramas and 12.1 channel spatial audio. Note also that a single computer equipped with 4 graphics cards may be able to drive 16 displays, but to do so efficiently a degree of parallelisation is almost certainly required, thereby ruling out a great deal of existing software and making the creation of new software a task similar in complexity to that of a cluster.

somewhat alleviated by compressing the images, or perhaps accommodated by a high-bandwidth networking technology like *Infiniband*.[154] However neither of these potential solutions can circumvent the second disadvantage, which is that the entire panoramic image must be generated by a single computer. Whether this be decompressing a stereo-panoramic video or rendering a real-time scene, the computational and rendering power of the master computer quickly becomes a bottleneck, restricting the resolution, quality and complexity of the imagery to 1/12th of that which could be generated by a single computer for a single display.

A different approach is to distribute the work of creating or decompressing images between all computers in the cluster. In the case of the AVIE, the most obvious division of labour is to have each render-node compute 1/6th of the panorama - the portion of the panorama illuminated by the projector(s) to which it is connected - and no more. For example, a panoramic video can be sub-divided and compressed into six video files, one for each render-node. Each render-node simply decodes, distorts, blends and displays their portion of the video, and apart from the problem of displaying each frame at exactly the same moment (a non-trivial problem in itself, discussed below), no communication between master and render-nodes is required whatsoever. A perfect case of divide and conquer.[155]

Such an approach, however, becomes somewhat more complicated with real-time simulations, in which the simulation cannot be so easily partitioned among render-nodes. Consider, for example, the simulation of a thousand balls, bouncing around within the interior of the AVIE. Each frame, 60 times per second, the states of the balls (positions, orientations, velocities etc) are calculated according to a simple emulation of real-world physics (collisions, aerodynamics, friction and gravity etc).

---

154 InfiniBand Architecture Specification Volume 1, Release 1.0, October 24, 2000, available from the InfiniBand Trade Association, *www.infinibandta.org*.

155 This approach is often used to display panoramic video in the AVIE. The panoramic image is divided into 60° sections, and compressed into 6 films, one for each render-node in the cluster. Blending and warping is performed in real-time, as the images are presented. When free rotation of the video is desired, the panorama is divided into 15° slices, and each render-node displays 4 or 5 slices. This approach is also a very effective means of multi-threading the video decompression, and very easily achieved with the icMPEG2 module.

Suppose also that a user is given interactive control of the direction of gravity's pull on the balls, using a simple joystick. Such a simulation would be by nature incremental, in that the positions of the balls are computed from their previous states.

There are now a number of ways of dividing the two tasks of computing the positions of the balls and rendering the images across the cluster. One on hand, the position of the balls could be computed each frame solely on the master, and their positions transmitted to the render-nodes, which they use to render their portion of the panorama. With this approach, the render-nodes need know nothing about the inputs of the user, nor the previous states of the simulation; they need only know the instantaneous position and orientation of the balls that are visible within their sector of the cylinder.

Such an approach may be viable with a thousand balls, but with a million, data transfer from master to render-nodes once again becomes a bottleneck. And suppose that the simulation is not rigid-body, but the balls deform on impact or as they spin through the air. Now, the position of *each vertex* of each ball must be distributed across the cluster.

An alternative strategy draws on the deterministic nature of the computer. The key observation here is that, assuming that all computers start in exactly the same state, and the simulation is entirely deterministic, and the time-step of the simulation is kept synchronised across the cluster, and the external inputs into simulation are duplicated on all computers, then the computers can be guaranteed, for any time in the future, to share the same picture of the virtual world. With such approach, all computers maintain a complete simulation of the virtual world, and it is only the (changes in) external inputs into the simulation that require distribution through the cluster.

The greatest challenge to such an approach is guaranteeing that the computers do in fact behave identically, for despite the formal determinism of computers, it is surprising how difficult this is in practise. The iterative nature of many interactive applications lends them a high degree of fragility, or brittleness, for if there is any difference in state between two machines, the iterative nature of the simulation will quickly expose it, sending the two machines on ever-diverging trajectories.

This strategy also requires a great deal of redundant computation. The physical simulation of the balls - detection of collisions and integration of their trajectories - is now being performed identically on all computers, and although each render-node need only render the portion of the world visible to itself, it must nonetheless maintain a complete simulation of the world, duplicating all the calculations made by its peers. Intriguingly, while redundancy is often used in engineering to provide a complicated system with robustness and resilience to error, here it is the very source of fragility.

These two approaches, the first where the simulation is performed on the master only and the instantaneous state of the world are propagated to the slaves, or the second in which each computer performs the simulation for themselves, represent two positions on a continuum. This can be seen by envisioning the entire AVIE system (both hardware and software) as a pipeline of deterministic data transforms. At one extreme are the input devices feeding external data to the master computer, while at the other, 12 projectors, connected to the 6 render-nodes, illuminate a screen. At some point between these two extremes, data must be distributed from the master to the render-nodes, and the key to designing effective software for a cluster is recognising at which point between these two extremes this distribution of data should take place.

Are there any approaches that might work in *all* cases, and therefore be made somewhat automatic?[156] One approach is to distribute only the inputs into the cluster: keystrokes, mouse movements, tracking data etc. However, suppose now that input into our simulation is not a joystick, but the output of a voxel-based tracking system. Viewers are invited to move about within the AVIE, and the centre-of-mass of their combined voxels controls the direction of a virtual gravity. In such a case, rather than distributing the raw voluminous input data to all computers, computing the gravity on the master alone is a clearly more elegant and efficient solution.

An alternative approach is to return to the other end of the pipeline, but rather than distribute the outputs of the GPU (the images), distribute the *inputs* into the GPUs. This is made possible by noting the layered structure of the graphics pipeline, in which all

[156] See Streit et al. (2004) for an overview of the various stages in the graphics pipeline that data might be distributed.

communication with the GPU takes place through a programming interface, such as OpenGL or DirectX. By using a modified 'cluster-aware' version of OpenGL, all communication to the GPU could be automatically divided and distributed across the cluster. Examples of this approach include *WireGL* (Humphreys et al., 2001), *Chromium* (Humphreys et al., 2002) and *Equalizer* (Eilemann et al., 2009). It is not difficult, however, to conjure scenarios in which such an approach fails. Suppose our simulation of the balls is being partially performed on the CPU and completed on the GPU, and every frame a great deal of data is passed from one to the other. This data must be now distributed across the cluster. Or consider applications such as *La Dispersion Du Fils* or *TVisionarium*, in which a great deal of video data must be asynchronously streamed and buffered to RAM from disk, operations that are outside the scope of a graphics API, but nonetheless must occur on the render-nodes.

In practice, the optimal distribution of data and computation is highly dependent on the application in question, and it is hoped that these examples have demonstrated this. It is for this reason the *Virtools*/iCinema SDK provides a flexible framework that permits and demands the application programmer to specify exactly which aspects of the simulation are computed on each computer, and which data to distribute each frame. The icCluster module provides tools for the creation of distributed objects – instances of *Virtools* classes (3D entities, meshes, materials, lights, as well as data-types vectors, matrices, arrays) that automatically broadcast any change in state across the network. icCluster also provides a clustered messaging and remote-procedure call framework, and the automatic distribution of input devices.

Asymmetry

When deciding upon a distribution strategy for a new project, the goal is to minimise four aspects: a) data transfer between computers b) inefficient or unbalanced use of computing resources (CPU, GPU, disk and RAM), c) duplicate computation, and d) complexity.

A principal source of complexity arises from *asymmetry*, which is a measure of the inhomogeneity in information about the virtual world/system available on each computer. In a completely symmetric design, all computers would have equal access to information about the state of the world, and this information would be the same on all computers. For example, suppose we wished to visualise, on the AVIE screen, the forces at work in the simulation of the balls. If the simulation were being calculated on all computers, then all machines would have access to this information. However, if the simulation were being run on the master only, and only the positions of each ball were being distributed to the render-nodes each frame, the render-nodes would not have access to this data, and would be unable to render it. The world model is asymmetric, and symmetry can only be restored by distributing the forces, and all other information produced by the simulation, across the cluster.

Asymmetry is also a measure of the inhomogeneity in the overall states of the computers: speed, fragmentation and space of RAM, VRAM or disks, GPU or CPU load, background processes, or input or output peripherals, for example. To distinguish between these two distinct forms of asymmetry, we might call the former *data-asymmetry* and the latter *system-asymmetry*. In practice, when using an operating system like Windows, a degree of system-asymmetry is unavoidable in a cluster.

In most AVIE applications, there is always a degree of asymmetry between the master and the render-nodes. For example, the master alone has direct access to the sound system. As spatial audio can be computationally expensive, it need only be computed on the master, introducing a considerable amount of system-asymmetry. Further, there is considerable data-asymmetry, unless an explicit effort is made to distribute the complete state of the sound-engine, such as the number and duration of sound-files, and the continual update of all their state parameters (position, volume, timing etc).

The master alone has direct access to input devices, introducing further asymmetry, and as the master is not connected to any projectors, it has no concept of a 'visible portion of the world,' [157] producing further asymmetry. Further, as render-nodes

[157] Sometimes the master is set to display nothing at all, save for a control interface, while at other times it might show some birds-eye view of the virtual world, including a representation of the AVIE and the viewers, functionality provided to all AVIE applications through icAVIEBase.

possess different outlooks on a virtual world, and so tend to render images of unequal complexity, a certain degree of system-asymmetry is always present among render-nodes.

In short, some degree of both data and system asymmetry is inescapable aspect of cluster software. This is important, for asymmetry has proven to be both *the principle source of design complexity* and *the principle source of design error* in software for the AVIE. To see how, let's return to our bouncing balls. Suppose we wish now to add sound to our simulation by playing, at random, one of ten sound-samples whenever a collision is detected. Sound is solely the responsibility of the master computer, but if the call to `rand()` made when selecting a sample is performed only on the master and not on the slaves, the system immediately enters a highly unpredictable and asymmetric state, for `rand()` is an iterative function, and all subsequent calls to `rand()` on the master will return a different value than on the slaves.

To keep the cluster in synchronicity, the same calls to `rand()`, or any other iterative or state-altering functions, must be made in the same order, at the same time, on all machines. This example is nothing other than programmer error, for the machines are doing exactly as they are commanded. Nonetheless, it highlights the fragility of maintaining an asymmetric world-model, in which certain data and processes are performed on only a subset of the machines, for the programmer must be aware of all the functions, methods and so on that might alter or depend upon some global state, like `rand()`. This particular frailty arises frequently when using 3rd party software, such as game-engines, where all such changes to inner-states are not always visible to the developer.

Another source of unpredictable behaviour is the use of multi-threading, where the non-deterministic advance of one thread relative to another can lead to temporarily uncertain states, breaking the requirement that the simulation be entirely deterministic. For example, the 'main-thread' of simulation may launch a 'worker-thread' to load from disk a sequence of images, and only once this action is performed, display them. However, the number of cycles of computation performed by the main-thread before the worker-thread completes is, from the point of view of the simulation, indeterminate, and likely to be different on different computers. This is also a good example of where system-asymmetry (e.g. heavy CPU load on one computer), can give rise to data-asymmetry (e.g. different frames of a video on different computers).

Such errors, where the states of cluster machines slowly diverge over time, or where the states evolves in an apparently nondeterministic manner, belong to the most difficult class of errors to identify and remedy.[158]

For mitigating these errors, icAVIEBase and icCluster include methods for early detection of state-divergence between computers. Every time-step, a package of data is passed from the master to the slaves where it is checked for coherence, and any difference is immediately flagged. The contents of this package includes such things as the number of calls to `rand()` or the position of virtual viewpoint, and data selected by the developer that might serve best as a 'canary in a mine.'

Synchronisation

Time, in a visualisation cluster, marches to the beat of the frame. Every 16.6 milliseconds a new image must be generated and delivered to the projectors, a persistent deadline that provides the single most important overarching factor in the design of the hardware and software. Further, the display of these images must be perfectly synchronised, in order that they might form a seamless panoramic image, a challenge that has posed a problem since Grimoin-Sanson used a ring of film projectors to display his flight above Paris in 1900, and continues to this day. With the AVIE, there are a number of aspects to synchronisation:

Data synchronisation: Data from the master to the render-nodes is distributed once each frame. This is typically a single, compressed, UDP packet. Once computation for the frame is complete, each render-node sends a single packet in reply.

Vertical synchronisation (VSync): New images are prepared in the back-buffer of a graphics card, while the projector vertically scans and displays the contents of the

[158] That a great deal of work was invested by the creators of *Virtools* in detecting, documenting or removing these potential sources of error was a key factor in adopting the software as a foundation for the iCinema SDK.

front-buffer. If the buffers are swapped while the projector is mid-way through scanning the front-buffer, then a mixture of the new and old image will be displayed, a visual artefact known as *tearing*. To prevent this, swapping of buffers is delayed until the vertical blanking interval, the small period of time between the end of one scan, and the commencement of the next. In this manner, VSync effects the coupling of the frame-rate of the renderer with the frame-rate of the projector.

Frame-sync/Genlock: The synchronisation vertical scan between computers. This is typically achieved by using graphics cards that permit the locking of the vertical scan to some external timer, or another graphics card.[159] There are, however, some software implementations (Allard et al., 2003; Waschbüsch et al., 2006).

Swap-sync: The synchronisation of buffer-swapping between computers. This can be achieved in software or hardware. With software swap-sync, each render-node sends a message to the master when it is ready to present a new image. Once the master has received such messages from all render-nodes, it broadcasts the command to go ahead. This communication takes time from the frame-budget, and is typically only accurate to within 1 or 2 ms. With hardware swap-sync, a direct electrical signal between GPUs provides a fast and precise barrier for swapping the buffers, removing the need for any network messaging and achieving a much higher temporal accuracy.

The AVIE uses different synchronisation strategies for the active frame-sequential and passive polar-stereo systems. Active-stereo, because it shows the left and right images in temporal sequence, and the shutter-glasses must be synchronised with the projected images, demands that all projectors maintain a very high-level of synchronicity. Here, hardware frame-sync is used, in concert with VSync. Hardware swap-sync is not, however, used. In the passive system, while temporal discrepancies might lead to tearing, they cannot result in one eye seeing an image intended for another, so the timing need not be so accurate. Therefore, software swap-locking is used and VSync is disabled, as the vertical retraces of the computers are unsynchronised. A key difference between the two is that active stereo systems must, therefore, use

[159] In the AVIE, *Quadro G-Sync* cards are used.

genlock-able 'workstation' graphics cards, while the passive stereo systems are free to use any graphics card whatsoever.

Conclusion

The grouping of multiple commodity PC computers into a single, synchronised computing cluster is a common strategy when creating multi-projector display systems. In adopting this strategy, however, the system is rendered incapable of executing existing software, save for a very small handful of 'cluster-compatible' applications, and the task of creating new software becomes significantly more costly and challenging. Finding methods for mitigating these aspects of the cluster-based approach to immersive display remains an interesting and open research challenge.

3.12 INTERACTION DEVICES

One of the features of the AVIE that distinguishes it from its panoramic predecessors is interactivity - the system senses and responds to the actions and movements of the user. This is achieved using a number of different interaction apparatus.

For many scenarios, a simple pointing device provides a suitable general purpose interface. Here, three degree-of-freedom orientation sensor was encased in a plastic case, designed to emulate a simple hand-held flashlight, and a number of buttons provided extra functionality. This device was often used to control the position of a virtual 3D cursor, and could be used much like a 3D mouse to select objects, "drag and drop", or steer the direction of motion. Sometimes this pointing device is used in concert with a joystick, mounted centrally within the AVIE in a custom designed console.[160] This console can be rapidly removed or replaced within the AVIE as needed.

More recently a small tablet computer has been adopted in place of the pointing device and console. The orientation sensors in the tablet allow it to function as a pointing device, while the touch-screen provides emulation of the console's joystick and

[160] The AVIE interaction podium was designed by iCinema colleague Volker Kuchelmeister, and constructed with the assistance of *Tiller Design*.

buttons. This approach provides greater functionality and flexibility than the console, as it allows for the design of dynamic, reactive, animated and context-sensitive interfaces tailored specifically for each application or experience.

Figure 56 - AVIE interaction console. iCinema & Tiller Design 2007.

3.12.1 Immersitrack Tracking System

None of these physical input devices, however, are able to track the position or movements of the users. For this, a much more elaborate vision-based tracking system was implemented. Developed by iCinema doctoral student Anuraag Sridhar, the *Immersitrack* tracking system uses an array of 16 cameras to follow the audience within the AVIE. To allow the system to work in near or complete darkness, and to avoid interference from the inconstant light of the projectors, the cameras operate in the near infra-red spectrum (approximately 830 nm). The cameras are fitted with infra-red filters and the AVIE arena is flooded with infra-red light, all invisible to the human eye.

Nine of the cameras are suspended above the AVIE arena, facing downwards, with the remaining seven cameras distributed around the top lip of the AVIE screen, providing oblique views of the AVIE environment. The images produced by the cameras are distributed among a small cluster of computers (four slaves and one master), where a variety of image processing and classification algorithms extract information about the position, movement, posture and gestures of the viewers.

At the lowest level, the system tracks the 2-D position of each user over time, and recognises simple events such as a user entering or exiting the system. Upon entering the AVIE, the system assigns the new user with a unique identifier, and uses a number of heuristics and predictive filters to maintain this identity over time. In addition to this, the system traces the 2-D contour of each user, as well as the 3-D position of the head, centre-of-mass and 3-D bounding box, all of which can be used for more nuanced spatial interaction. The system also applies noise reduction filters to remove any jitter and uncertainty. See Sridhar and Sowmya (2011) for details.

Higher level capabilities of the *Immersitrack* system include the real-time construction of 3-D voxel models of the user (Sridhar & Sowmya, 2009), as well as a pointing-gesture tracking algorithm, allowing people to use their fingers as interaction devices (Sridhar & Sowmya, 2008).

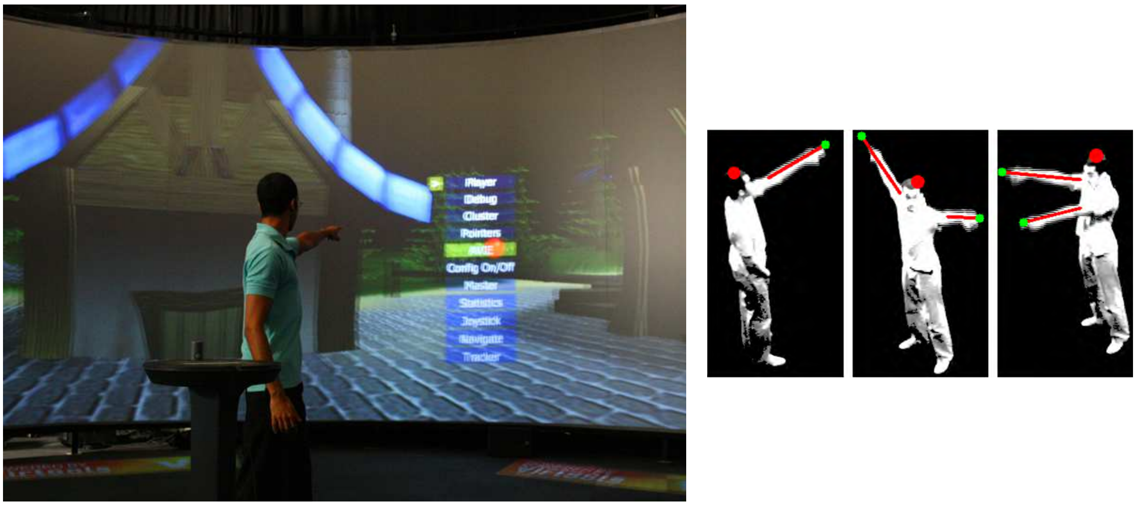


Figure 57 - The Immersitrack finger tracking system permits a user to manipulate virtual objects or graphical interfaces with pointing gestures. Images from Sridhar and Sowmya (2008).

The system can also detect and recognise simple gestures and activities, using an approach to expert-system classification known as *Ripple Down Rules* (Compton et al., 1991; Sridhar et al., 2010). The recognition system has been used to identify simple actions such as crouching, standing, jumping, running or walking, as well as simple arm postures. All of these features of the tracking system are described in great technical detail in (Sridhar, 2012).

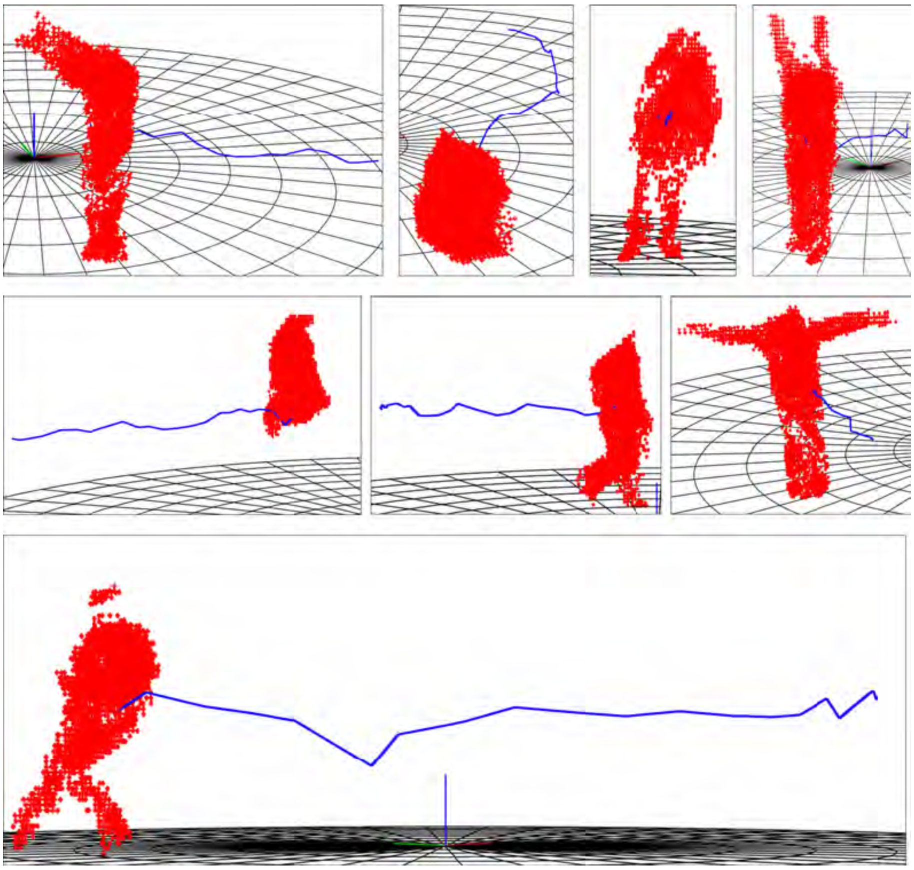


Figure 58 - Immersitrack gesture recognition system can recognise such actions as crouching, standing, jumping, walking,running or arms outstretched up, forward or to the side. Image reproduced from Sridhar (2012).

3.12.2 Using the tracking system

Information about the state of users within the AVIE, including such things as their centre-of-mass and bounding box, is fed continually from the *Immersitrack* system to the AVIE system through the *icAVIEtrack* module of the iCinema SDK. The *icAVIEtrack* module maintains a simple object-oriented model of all tracked users which the application developer accesses via a simple API (in C++ and VSL). In addition, *icAVIEtrack* provides a visual debugger, embedding graphical 3D representations of the users within the virtual world. This allows the state of the virtual users to be monitored on displays outside of the AVIE, or within the AVIE itself. The module also allows for the recording and playback of user movements, as well as a graphical user-interface for manually simulating the movements of people within the AVIE using a mouse or keyboard, two tools that prove indispensible for development and debugging.

An important aspect of the model maintained by *icAVIETrack* is that there is no distinction made whatsoever between real entities (in this case, the users, although the tracking system is capable of tracking other objects) and virtual entities. *icAVIETrack* shares its spatial model with *icScenarioManager*, the module responsible for reasoning about "spatial narratives" by continually detecting and classifying spatial events and relationships between entities, such as collisions and separations, accelerations and trajectories, approaches and departures, entrances and exits and formation and splintering of groups. As both classes of entity, real and virtual, are treated identically by the *icScenarioManager*, such spatial events are detected both in the virtual and real worlds, and between the two. The resulting model and detected events greatly facilitate the construction interactive 'co-evolutionary' narratives, in which narrative events take place in both the real and virtual worlds, and at threshold of the two. For example, in Figure 59 below, an autonomous virtual agent can be seen pursuing a real user, represented here by their real-time voxel reconstruction. The virtual character is programmed to look at, and then approach, any nearby moving object, be it virtual or real, a behaviour easily achieved thanks to the seamless union of the *Immersitrack* system and the *icAVIETrack* and *icScenarioManager* modules.

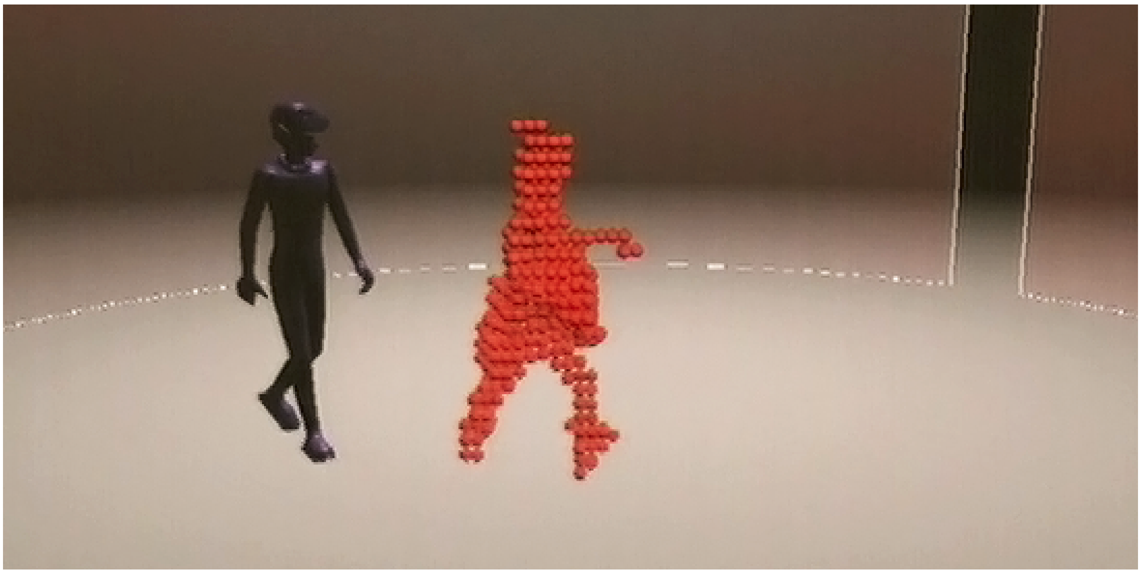


Figure 59 - Virtual/Real agent interaction. A virtual autononomous character follows a voxel representation of a real user around the AVIE.

This approach to fusing the real with the virtual is used to great effect in the iCinema project *Scenario* (Del Favero & Barker, 2010), in which the viewers participate in and influence the unfolding narrative by physical interacting with virtual characters and objects.

Figure 60 - Virtual and real characters interact through physical motion. Here a player engages a virtual character by physically approaching it. (Scenario, iCinema 2010).

3.13 THE iCINEMA SOFTWARE DEVELOPMENT KIT

The AVIE is designed to serve as a multi-purpose VR theatre, capable of immersing an audience in a variety of different experiences. The computer programs that enables these experiences tend to share a great deal of common features and structures, regardless of how different the perceived experiences may be. For example, mechanisms for rectifying projection distortion, spatialising audio, rendering stereo panorama, distributing data amongst machines or handling input devices, are common to all AVIE applications. There should be no need to re-invent or re-implement tools and techniques with each new project or experiment. The *iCinema Software Development Kit (SDK)* was created to accelerate development of AVIE applications by encapsulating and abstracting all that is common and necessary for the platform. This software is very briefly described in this section.

3.13.1 Game Engines and VR Toolkits

The development of platforms - collections of re-usable components with layers of abstraction - is a foundation stone of both software and hardware computer engineering. Re-usable components help avoid a constant re-inventing of the wheel, while abstractions permit the creation of higher-level functionality, given limited development resources and time. In creating a platform, however, there is always a certain compromise between versatility and simplicity. When creating a platform for a virtual reality system, one is faced with finding a balance between these two poles.

One approach to creating a VR platform is to adopt and adapt a computer game engine for use in a VR system (for example *CaveUT* (Jacobson et al., 2005; Jacobson & Lewis, 2005), *CryVE* (Juarez et al., 2010), *BlenderCAVE* (Gascón et al., 2011), *CaveUDK* (Lugrin et al., 2012), or *getReal3D* and *MiddleVR*[161] for the *Unity* platform). This approach has a number of advantages. First, computer games have come to represent the forefront of real-time interactive software, and have become the standard by which these interactive experiences are judged. Game engines integrate many features and techniques frequently called upon in VR applications, such as real-time graphic rendering systems, animation systems and physics simulations, asset management and 3D sound engines, so by using a commercial game engine, one gains access to the latest developments in all of these different technologies. Second, the task of keeping up with the unrelentingly evolution in computing hardware and software is delegated to the developers of the game engine, who, armed with budgets measured in millions (and profits in hundreds of millions), are somewhat better equipped to keep abreast than most. Third, by using a commercial platform, one gains access to a community of developers and repository of knowledge that might not be so readily available with more bespoke software. Finally, the software is subject to a high degree of quality control.

---

[161] MiddleVR is an attempt to extend the *Unity* game engine with VR capablities, developed by Sebastien Kuntz, a former developer of the Virtools VR modules. See *www.imin.fr/middlevr-for-unity.* Similarly, *getReal3D* is a Mechdyne's VR plug-in for *Unity*. See *www.mechdyne.com*.

However, with these benefits come significant disadvantages. For one, almost all game engines require modification or extension before they can be used on multi-computer, multi-projector display systems. The most significant of these modifications are those required to enable parallel execution across a cluster of computers, but also include such things as support for unconventional sound systems, or unconventional rendering algorithms. Although most game engines provide some manner of extensibility, from plug-ins to scripting languages, very often these modifications require changes or additions to the core of the framework (see for example (Lugrin et al., 2012)). Even in cases where the relevant source-code can be acquired (which is very often not the case, given the highly competitive nature of the game industry), modifying existent software as large and complex as a game engine is a significant challenge, requiring an intimate knowledge of the kind typically only enjoyed by the creators of the software themselves. (And without access to budgets that more typical customers might enjoy, it can be very difficult to gain the attention of these creators.) The typical outcome is a version of a game engine that is in some way crippled or reduced, where a subset of features remain incompatible with the VR platform.

Second, the benefits of a game engine - ease of use or graphical realism - are gained at the expensive of versatility. A game engine incorporates a vast multitude of optimisations and simplifications based on assumptions about the kinds of experiences - that is, games - that will be created. For example, an animation system might assume bipedal characters, an occlusion system might be based on simulation of rooms and corridors, the text or video system might be limited to two-dimensional "heads-up" display, or the geometry system might assume that all "assets" are created by artists, and so have limited support for generative or procedural modelling. The result is that as the software is used for less and less "game-like" experiences, more and more of the features of the game engine become irrelevant. Or worse, the structures imposed by the game engine become impediments.

Third, the high quality attained in computer games is not solely due to the game engine, but more a result of the extraordinary amount of man-power, from both programmers and artists, that goes into creating these games. In this respect, a sophisticated game engine, having been designed with these large teams of artists and developers in mind, may produce less than sophisticated results in the hands of a small research team. Finally, given the amount of time and resources required to adapt a game engine for use in VR, and the time required to become fluent with said engine, adopting a commercial game engine as the basis of a VR platform is to enter into a long term relationship with a third party, and place at the heart of one's platform an element whose evolution and longevity is largely beyond one's control.

An alternative to a game engine is to use one of the very small number of commercially available virtual reality authoring systems, such as *WorldViz* or the *EON Reality* software suite.[162] In contrast to game engines, these tools natively support common virtual reality platforms (such as the CAVE or head-mounted displays). However, these products are predominately intended for the engineering industries where VR is used for prototype visualisation, and seldom offer more than simple visualisation and manipulation of static models. They tend to be expensive and offer less flexibility and extensibility than a typical game engine.

Another approach is to adopt one of the numerous open-source virtual reality software development kits created by VR research institutions in response to the very same problem. Examples include *Avocado* (Tramberend, 1999), *VRJuggler* (Allen Bierbaum et al., 2001; Aron Bierbaum et al., 2005), *Syzergy* (Schaeffer & Goudeseune, 2003), *FlowVR* (Allard et al., 2005; Allard et al., 2010), *CalVR* (Schulze et al., 2013) and *FreeVR* (Sherman et al., 2013). The principal advantage of these libraries is that they are designed almost entirely around the problems specific to VR. Cluster-distribution, multi-projector display and 3D input devices, for example, are cornerstones upon which the software is structured. Beyond this, however, they tend to provide little more than proofs of concept. They live principally as research platforms, and are seldom used to create anything as rich or complex as might be found in a computer game. They tend to suffer from very limited user bases, and without commercial support to finance dedicated developers, enjoy only sporadic development or maintenance. In summary,

[162] See *www.worldviz.com* and *www.eonreality.com.*

these toolkits may serve as good starting points, but require considerable extension if they are to be used to create rich and compelling virtual experiences.

Another alternative is to build an entirely new VR platform oneself, using a variety of existing components and libraries (otherwise known as 'middleware'). By combining software libraries devoted to particular aspects of an interactive immersive application, such as sound (e.g. *OpenAL*), graphics (e.g. *OpenGL*, *Ogre*), interaction (*VRPN*, *OpenNI*), physics simulation (e.g. *ODE*), parallel computing and inter-computer communication (*MPI*, *ZeroMQ*), a bespoke development platform can be tailored to exactly meet the needs of the developers. For certain projects this might prove the most efficacious path, provided the project is suitably limited in scope, and sufficiently skilled software engineers are at hand.

In summary, in seeking a software solution for the system, there is no panacea, as all paths possess significant shortcomings.

3.13.2 *Virtools*

The iCinema SDK is built upon the *Virtools Dev* platform, a commercial product that sits somewhere between a game engine and a general purpose VR development platform.[163] Originally designed for rapid prototyping of game-like applications, *Virtools* was also intended for use by virtual reality developers. As a consequence, *Virtools* provided native support for cluster parallelisation, and although the mechanisms provided for parallelisation would require significant extension for use in large scale productions, a great deal of engineering was invested by the developers of *Virtools* to ensure their software was sympathetic to parallelisation. That is, they had ensured that there was nothing within the software that would prohibit its use in a distributed manner, and with *Virtools* being used in this way by other VR developers, a great deal of uncertainty was removed from the equation.

---

[163] Original called *NeMo* in 1996, and then *Virtools Dev* in 1999, the name was changed once more to *3DVIA Virtools* after being acquired by Dassault Systèmes in 2005. The software was discontinued in 2013.

Equally important in the decision to adopt *Virtools* was the design of the software. *Virtools* lacked many features that might be found in modern game engines, and what features it did provide were suitable only for rapid prototyping and not of sufficient quality or efficiency for production-quality projects.[164] However, *Virtools* was elegantly structured, highly modular, and highly extensible, providing many different avenues for extension. As such, it would serve as a suitable scaffold for the iCinema SDK.

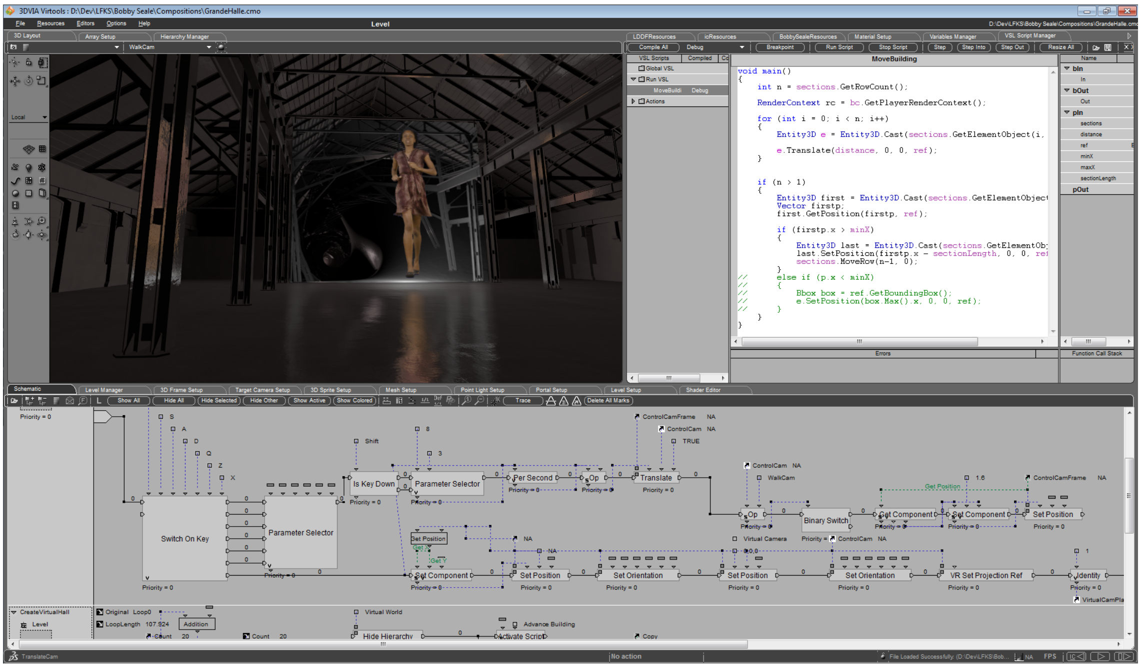


Figure 61 - Virtools integrated development environment.
In the lower window a portion of a script composed in the visual graph-based scripting language can be seen, and in the upper right the text-based Virtools Scripting Language *is visible. In the upper left is a 'live' view of the scene, in this case, a view of the* LFKs *project* Bobby Seale Got His 9 *VitaNONnova #3.*

Virtools takes the form of an integrated development environment, within which the developer creates 'compositions'. Compositions can be built using the *Virtools* graph-based visual scripting language, the *Virtools* textual scripting language (a subset of C++), with the C++ API, or a combination of all three. As such, *Virtools* is accessible to a broad range of developers with varying background and expertise, from students to scientists, or from artists to seasoned software engineers. The intention was that, by

[164] To give a concrete example, Virtools provided some very elementary shadowing mechanisms, none of which would be applicable to a production grade project. However, there was nothing within Virtools to prohibit the developer implementing his or her own shadow system.

basing the iCinema SDK on such a flexible and accessible tool, the AVIE would be opened to greater use.

3.13.3 The iCinema SDK

The iCinema SDK is a library of C++ methods and classes, *Virtools* 'plug-ins', compositions and scripts that, taken together, provide an integrated development environment for constructing AVIE applications. It is composed of modules that provide methods for running clusters of computers, warping and blending seamlessly images from multiple projectors, multi-channel 3D audio, a 3D graphical user-interface, shader management, high-bandwidth video playback, speech recognition, embedding web content, character animation and artificial intelligence, interfacing with input devices and tracking systems, and more. It provides a collection of base classes and foundational elements such as logging and debugging and configuration management, and provides a template for the structuring of new projects.

A key goal in the design of the iCinema SDK was that *any Virtools* composition could be easily, if not instantaneously, adapted to run in the AVIE. The would-be AVIE developer need only create a *Virtools* composition, without intimate knowledge of the workings of the AVIE system itself. In practice, the amount of adaptation required for a *Virtools* composition to execute on the AVIE is very much a function of the complexity of the composition. Architectural walk-through applications, where interaction is limited to simple navigation, are a perfect example of an application that can be instantaneously executed on the AVIE without modification.

The SDK is designed in such a way that compositions written for one VR platform can be easily migrated to another. For example, an application intended for the AVIE can be experienced on the iCinema iDome,[165] a single-user hemi-spherical platform also developed at iCinema (Kuchelmeister et al., 2009; Kenderdine, 2010).

The SDK features are accessible to a developer through C++ or the *Virtools* scripting language VSL. The SDK includes documentation and around 150 example compositions

[165] See *icinema.unsw.edu.au/technologies/idome*

demonstrating different elements of the platform. Comprising around 250,000 lines of code, a detailed description of the SDK cannot be given here. Instead, an overview of the various modules of the SDK can be found in Appendix G.

The iCinema SDK has been successfully used to implement a number of productions, including *Scenario* (Del Favero & Barker, 2010; Del Favero et al., 2010),[166] *Spaces of Mnajdra* (Flynn, 2012), *TVisionarium II* (Brown, Del Favero, McGinity, Shaw, Weibel 2008) (Bennett, 2008),[167] *Hampi-Live* (Shaw & Kenderdine, 2006; Kenderdine et al., 2007), *Rhizome of the Western Han* (Kenderdine et al., 2011) ,[168] *Bobby Seale Got His 9* VitaNONnova #3 (Bruyère & LFKs 2013) the *iCASTS* mining training simulators (iCinema & UNSW Mining Engineering),[169] *Pure Land: Inside the Mogao Grottoes at Dunhuang* (Kenderdine & Shaw, 2012),[170] and numerous AVIE demonstrations and experiments (Aymerich-Franch, 2010, 2012). It was also used by the author to teach the iCinema Studio class in 2009, in which students with no prior experience working with virtual reality or *Virtools* were guided through the process of creating applications for the AVIE.

Figure 62 - AVIE displaying 360° digital video captured with the iCinema SphereCam.

[166] *icinema.unsw.edu.au/projects/scenario*
[167] *icinema.unsw.edu.au/projects/t_visionarium/t_visionarium-ii*
[168] *alive.scm.cityu.edu.hk/projects/alive/rhizome-of-the-western-han-2010*
[169] *icinema.unsw.edu.au/projects/icasts---mining-vr*
[170] *alive.scm.cityu.edu.hk/projects/alive/pure-land-inside-the-mogao-grottoes-at-dunhuang-2012*

Figure 63 - TVisionarium II (Brown, Del Favero, McGinity, Shaw, Weibel, 2008).

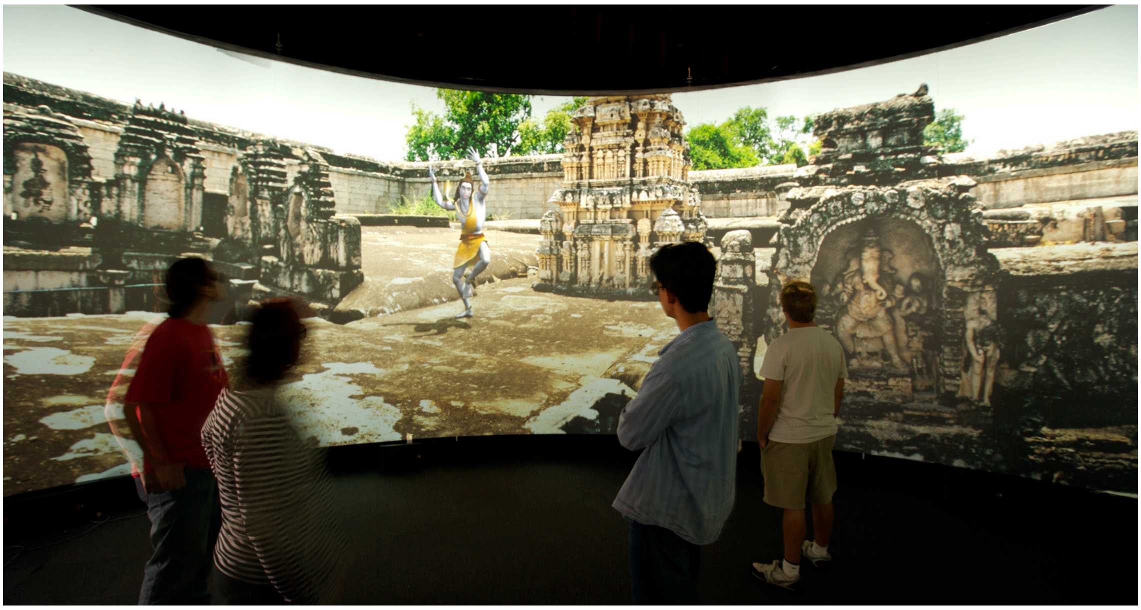


Figure 64 - Stereo panoramic video in Place-Hampi (Shaw & Kenderdine, 2006)

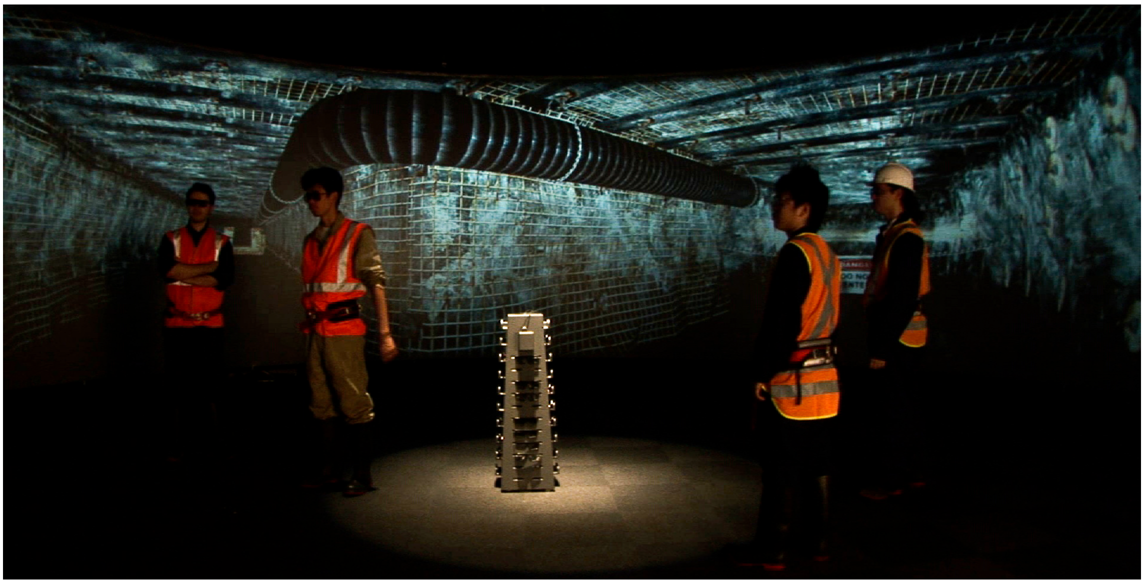


Figure 65 - iCASTS Mining Training Simulation (iCinema & UNSW School of Mining Engineering, 2009)

3.14 Analysis and Discussion

In this section, various strengths and weaknesses of the system are briefly discussed, and possible remedies proposed.

Resolution and Aliasing

As mentioned, the panoramic resolution of the AVIE is approximately 7570 pixels in circumference and between 900 to 1000 pixels in height. The pixel resolution for a viewer standing in the centre is, therefore, 21 pixels/degree or 2.85 arc-minutes/pixel. A viewer with 20/20 vision is able to distinguish lines 1 arc-minute apart, or 120 pixels/degree, so the AVIE display is by no means above the 20/20 threshold of human acuity. This single threshold is, however, a very crude measure of visual acuity, for the resolving power of human vision varies with a great deal of factors, such as brightness, contrast and shape, colour, motion and periphery of vision (Fulton, 2005). Under low light, visual acuity can be as low as 8 to 18 pixels/degree (Cowan, 2002).

In practice, the adequacy of the AVIE's resolution is almost entirely dependent on the type of images on display. Experience developing applications for the AVIE has demonstrated this resolution to be sufficient for a wide variety of imagery, from polygonal 3D graphics to panoramic video and photography. This is to say that the pixel artefacts pose little impediment to presence.

There are, however, particular visual elements and imagery that pose a challenge to the system, namely sharp, straight vertical and horizontal lines and small text. Such elements arise most commonly when attempting to use the AVIE as a conventional 2D display, as if it were a giant two-dimensional desktop. In such situations, it is not so much the resolution of system that must be taken into account when designing content, but rather the non-linear distorted manner in which the projector's pixels fall on the cylindrical screen. As seen previously, in order to display a straight line in the AVIE, a curved line must be projected. Consequently, it is impossible to render straight lines of any orientation (that is, straight in screen-space) with exactly a single pixels width, and some aliasing must occur. Anti-aliasing helps to reduce this display artefact considerably, and AVIE imagery is typically rendered with x4 or x8 multi-sample anti-aliasing. Nonetheless, it is not possible to display vertical and horizontal lines with the sharpness that is enjoyed on a typical desktop display.

This lack of single-pixel accuracy is even more pronounced when it comes to the display of fine text. One reason for this is that we have grown used to the clarity of text attained using *sub-pixel rendering*, a technique now common to all modern computers and LCD displays. By treating the red, green and blue sub-pixels of a pixel as picture elements in their own right, sub-pixel rendering effectively triples the resolution of a display. (Colour accuracy has been traded for spatial accuracy, and if one were to inspect text closely under magnification, one would see that the contour of black text is in fact multi-coloured). Projectors of the type used in the AVIE, however, do not use sub-pixels to create colour, but rather rely on a rapidly spinning coloured wheel to project a series red, green and blue monochromatic images. As a consequence, text cannot be displayed with the crispness and clarity typically enjoyed on a desktop computer.

Brightness

As mentioned in the section on stereoscopy, light efficiency is a major challenge for all forms of stereo projection, and the AVIE unfortunately is no exception. The area illuminated by one projector is approximately 1/6th of the screen surface, or 21m^2. The polar-light 'passive' AVIE systems use twelve *ProjectionDesign F20* projectors, each delivering 3300 lumens of light, but as a typical linear-polarising system has a light efficiency of just 32%, the effective brightness of each projector is only 1050 lumens. In addition, only 90% of the projected light falls on the screen, and 10% of this is in blending regions, reducing brightness further to 900 lumens. Assuming a screen gain of 2, this is approximately 2 x 900 / 226 = 8 foot-lamberts = 27.2 cd/m^2 (nits).

The results for the active-stereo systems however, are markedly different. A single *ProjectionDesign F10 AS3D* projector produces just 2000 lumens, and after light is lost to active-stereo (17%) and pixel-efficiency (86%), all that remains is a meagre 2000 x 0.17 x 0.86 = 292 lumens. With a white Lambertian screen material with a gain factor of 1, this results in just 1.3 foot-lamberts (4.45 cd/m^2). If a silver high gain screen is used this figure can be doubled to 2.6 foot-lamberts, but note that a high-gain screen may introduce an unwanted polarisation of light, which would result in a diminished and uneven brightness due to the active-shutter glasses being also polarising filters.

How much light is required? The Society of Motion Picture and Television Engineers (SMPTE) standard for 2-D cinema is 14 foot-lamberts.[171] In practice this level of brightness is almost never attained for 3D cinema, and levels as low as 6 foot-lamberts or even 4.5 foot-lamberts are not uncommon in today's cinemas.[172]

The critical threshold is 3 foot-lamberts. As luminance drops below 3 foot-lamberts, photopic vision (day vision) starts to give way to scotopic vision (night vision). As night falls, vision relies less on the eye's cone photoreceptors, which provide colour vision and high acuity, and more on the eye's rod photoreceptors, which provide high light sensitivity, but relatively poor acuity and no colour vision. Between pure photopic and scotopic vision there is a transitional stage, known as *mesopic vision*. The properties of mesopic vision are complex and intricate (Stockman & Sharpe, 2006), including changes in spatial, spectral and temporal acuity and a shift in colour perception known as the *Purkinje effect* (Purkinje, 1825): as light levels drop below 3 foot-lamberts, not only does colour vision diminish, but colours are distorted, with a loss of sensitivity to the red end of the spectrum, a shift in perceived brightness towards the blue spectrum and an overall loss of colour vibrancy.

Such colour distortions have been observed in the first generation of active-stereo AVIE configurations, which is unsurprising given the 1.3 foot-lamberts is well within the mesopic range. Unfortunately, any prospect of simply adjusting the colour of the projected imagery to counter these distortions is complicated by the fact that adaptation to low light is a slow, gradual process, taking as long as 20-30 minutes. Perception of colour in the active-stereo AVIE depends, therefore, on how long one has been in the theatre. More than this, the author has observed a great deal of variation between

[171] Nominal luminance of peak white specified for digital cinema is 48 cd/m^2, as specified in SMPTE (2006).

[172] Achieving sufficient brightness is a major hurdle for 3D cinema. According to R. Ebert (2010), "the vast majority of theaters show 3-D at between three and six foot-lamberts (fLs)."

individuals, with some reporting an almost complete lack of colour perception, while others seem to perceive only very subtle shift in colour.

This problem is remedied in the more recent active-stereo AVIE installations which use five active-stereo 7500 lumen *ProjectionDesign F35 AS3D* projectors, bringing the screen brightness closer to 4 foot-lamberts.

Black levels, Inter-reflections and Contrast

Somewhat paradoxically, the AVIE projectors also suffer the problem of producing *too much* light. The DLP projectors are incapable of creating a true black, for they tend to 'leak' light. This has three undesirable consequences. First, even when projecting a completely black image, the combined light from the 12 projectors is still sufficient to light the theatre. It is impossible, therefore, to plunge the audience into a true darkness in which they are unable to perceive themselves or on another. Second, the raised black levels reduce image contrast, and third, the raised black levels lead to visible blend regions whenever black is projected; the non-black black is twice as bright in the overlap regions. This last issue can be remedied by raising the brightness of the black projected between blend regions to match the doubled brightness of the blend regions. However, this is done at the expense of overall image contrast and is seldom considered a desirable compromise.

An unavoidable consequence of the cylindrical screen is the reflection of light from one part of the screen to another. Such secondary reflections result in an overall diminishing of contrast, and are particularly pronounced when bright imagery is displayed, such as a scene with a white background. Such inter-reflections are unavoidable, and must be accommodated when designing content for the system, largely by avoiding large white objects or backgrounds.

There are, however, certain situations where this inter-reflected light can add realism to the scene. Consider, for example, an indoor scene with a fireplace or flickering television as the sole source of light. This light will illuminate not only the opposite sides of the room - a very crude but sometimes effective real-world approximation of global illumination - but also illuminates the audience, helping to embed them within the virtual scene.

Stereo Contrast

The active-stereo systems provide a very high level of stereo-contrast. For most of the field of vision, there is no perceptible stereo-ghosting whatsoever, and it is only in the periphery of vision that very subtle stereo-ghosting can be discerned, and then only when sharp high-contrast imagery is displayed. This drop in stereo-contrast in the periphery of vision can be attributed to the effectiveness of the shutter glasses varying with the angle that light strikes the glasses.

With the polarised-light passive system, however, ghosting is significant. While at the top of the screen ghosting is largely imperceptible, at the bottom of the screen, where angles of incidence and reflection are highest, loss of stereo channel separation is almost complete, confirming the predictions presented in Section 3.5.6.

Uniformity of image

Despite the wide range of angles of incidence and view, the image is surprisingly uniform over the entire screen. Only under certain conditions (when bright monochrome test images are displayed) can a subtle brighter ring be discerned about 1/3 from the top of the screen, and when projecting normal imagery this is imperceptible. Most importantly, there is no perceptible change in image brightness, colour, contrast or stereo contrast when the viewer moves within the theatre.

Shadows and Glare

Shadows cast by audience members onto the screen, when they approach the screen, prove to be insignificant. A viewer, 1.7m in height, begins to cast a shadow at a distance of 2.3 metres from the screen. However, the angle of projection is sufficiently high that even as one's shadow grows (at 2m, our 1.7m tall viewer casts a shadow 33cm high), it remains on the periphery of view.[173] Once the viewer is sufficiently close to the

[173] Shadow height $s = (hR - Hr)/(R - r)$, where R is distance of projector to screen (6.25m), r the distance of viewer to screen, H the projector height (4.6m) and h the viewer height (1.7m).

screen in order to cast a significant shadow (1m distance for a ~1m shadow), the resolution of the image is insufficient to maintain the viewer's gaze. It is, therefore, not a situation that the viewers tend to find themselves in.

However, it is observed that people tend to approach the screen *backwards*, facing the opposite direction, seeking a view of the entire cylinder. In this case, it is not shadows that are a distraction, but the *direct projection of light into the eyes*. As such, the "shadow-free zone" exists, but should rather be called the "glare-free zone". For an eye-height of 1.6m, this glare-free zone is a circle with radius of 2.83m. In practice, neither glare nor shadows have presented an obstacle in the enjoyment of AVIE experiences.

Depth-of-focus

Depth of focus is sufficient to provide sharp image over the entire image surface.

Field of View

The immersiveness provided by the complete 360° panoramic view of the AVIE screen is compromised somewhat by its comparatively limited vertical field-of-view. For the central viewer, the 4m high screen fills just 44° of their vertical vision – approximately one third of their natural vertical range.[174] Of all the shortcomings of the AVIE, this is perhaps the most significant, for a number of reasons.

First, it greatly restricts the display of entities and structures inside the AVIE, for unless they fit somehow to within the (vertically) narrow frustum of view, they will be cropped by the top or bottom of the image frame. Any virtual entity standing on the ground will necessarily be cropped by the bottom edge of the screen the moment it enters the screen perimeter. As can be seen in Figure 66, this largely prohibits virtual characters from entering the AVIE space, imposing a form of barrier between the virtual and real worlds. This barrier is most visible when designing AVIE experiences that aim to provide the audience with some form of physical causal agency over the virtual

[174] To completely fill their vision, the viewer would need to stand within half a metre of the screen, a distance where, resolution, divergent disparity, vergence-accommodation conflict and shadows would all prove problematic.

world, for without physical proximity to the virtual objects, interaction must be based on some form of "action at distance."

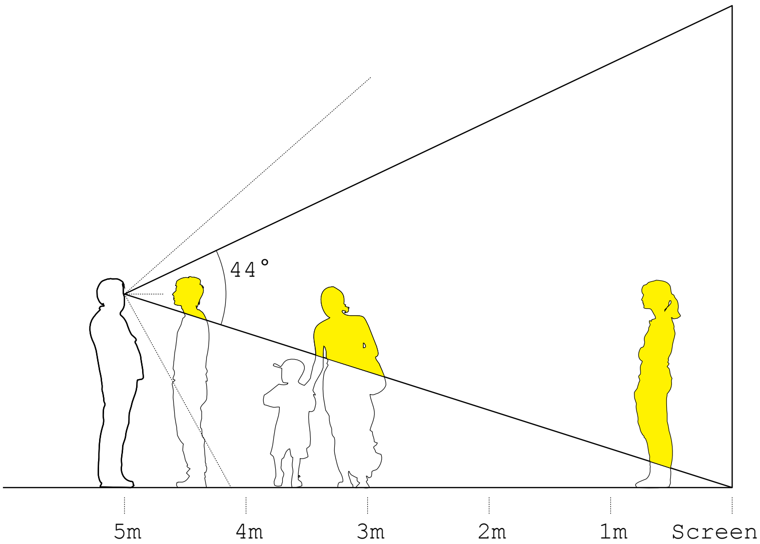


Figure 66 - Vertical Field of View of AVIE screen. Only the yellow portion of the virtual characters are visible.The dotted lines show typical vertical field-of-view of human vision.

Second, the visibility of the boundaries of the screen betrays the presence of a screen, although this has proved to be far less disruptive than previously imagined.

Third, the ground, in ordinary vision, very much dominates the visual field, providing strong visual cues for perception of self-location and self-motion. This is perhaps reflected in the asymmetry of the vertical field-of-view, which extends 60° upward, and 75° downward (Spector, 1990). When a stationary virtual viewpoint is adopted, the visibility of the real-world floor or ceiling may certainly appear odd, but it does not necessarily give rise to competing rest frames. When the virtual viewpoint is put in motion, however, the viewer is presented with two potential rest frames – the real-world floor of the AVIE and the virtual world. Here, black, unreflective and untextured carpet is used to reduce the visibility of the real-world floor, greatly increasing the likelihood that the virtual world will be perceived as a rest frame. When this occurs, the floor beneath the viewer is perceived as moving with the viewer through space - a strange floating black disc conveying the viewer through the virtual world.

Note that while it is certainly feasible to project imagery on the floor of the AVIE by adding more projectors, such an approach is incompatible with the omnistereo projection, for the omnistereo projection is predicated on the direction of the viewer's gaze being everywhere more or less perpendicular to the screen surface, *and* that the viewer's eyes are displaced horizontally to the left and right of this direction. In other words, it must be possible to infer with reasonable accuracy the orientation of the viewer's ahead from the orientation of the screen surface. With images projected on the floor and an audience free to move at will, there can be no such inference about the viewer's direction of view. It is for this reason that images cannot be projected onto the floor of the AVIE.

Vergence-Accommodation Conflict

As mentioned in Section 2.5.2, all stereo images present conflicting depth cues: for while the eyes converge at the distance of the virtual object, they must continue to focus (accommodate) on the image plane. This vergence-accommodation conflict can lead to discomfort and fatigue. Shibata et al. (2011) calculate a "zone of comfort" for viewing stereo images, where a range of acceptable vergence distances is given as a function of focal distance.

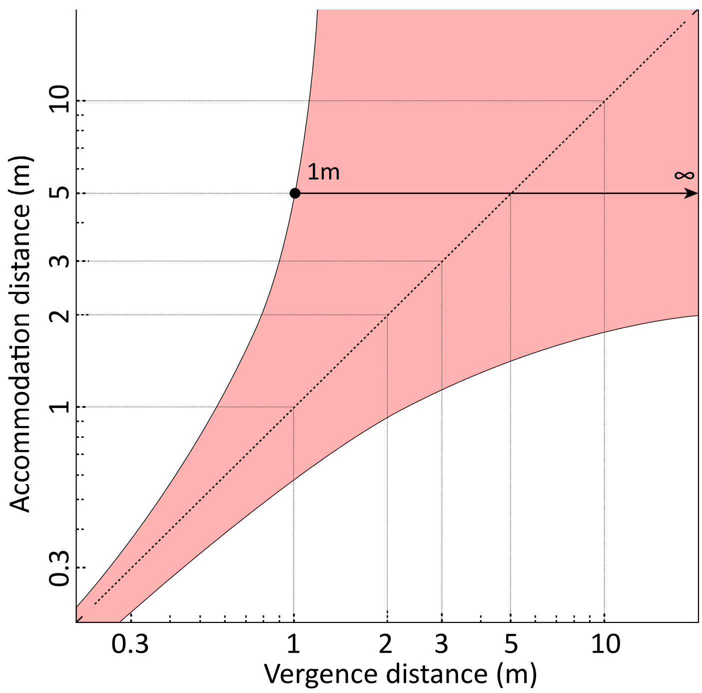


Figure 67 - Zone of Comfort: Vergence-Accommodation Conflict. Reproduced from Shibata et al. (2011).

For a viewer standing at the centre of the AVIE, with an accommodation distance fixed at 5m, it is suggested that virtual objects not be brought any closer than 1m from the viewer, and may extend to infinity. Indeed, for most of the AVIE, where the viewer is between 2m and 10m away from the screen, vergence-accommodation conflict does not present a problem.

Omnistereo

Overall, omnistereo can be considered as an effective approach to providing multiple free-moving viewers with an immersive experience. When omnistereo imagery is shown, the walls of the cinema quite tangibly disappear, and the audience all enjoy an equally valid view of the world, regardless of their direction of view. The depth cues of the projected imagery far outweigh the depth cues of the screen surface providing a tangible sensation that the AVIE has in fact disappeared and a world now stretches out beyond the bounds of the room. In this respect, the concept of omnistereo seems valid.

Although the geometry of the image is strictly only correct for a viewer standing at the centre of projection (and even then, not exactly), distortion has largely proved inconsequential to the experience of the AVIE. Here, it is the continuous curvature of the cylindrical screen that separates the AVIE from faceted display systems such as the CAVE. While omnistereo projection can also be implemented in a CAVE, allowing multiple viewers to view the world with different directions of view (e.g. (Naemura et al., 1998)), the corners of the CAVE's cubic screen present significant distortions for the non-central viewer.

There is, however, a marked difference between distortion observed by a moving observer and a static observer. For the static observer, distortion is essentially imperceptible for large parts of the AVIE. For the moving observer, however, the world tends to visible compress, dilate and skew as the viewer moves about, and these distortions grow in strength the closer they approach the screen. When the viewer moves about with the AVIE, their visual system accounts for the apparent lack of motion parallax by perceiving either the foreground, background, or both, as in motion, and as a result, the scene appears to 'swim.'

The presence of other audience members tends to enhance the perception of space, and immersion, by providing natural parallax and occlusion cues, as well as a measure by which to judge size and distance. This effect can be, at times, very powerful, as if one's fellow audience members are truly "within" the virtual scenery, and suggests that the use of real-world objects and 'props' in concert with the virtual imagery may be one path to heightened immersion.

Designing for the AVIE

There are, therefore, a number of features of the AVIE's projected image that suffer in comparison to a conventional single-user, single-screen display. Namely, the peculiarities of the omnistereo projection, the limited contrast and brightness of the image and, in comparison to other immersive systems such as the CAVE or a head-mounted display, a comparatively limited vertical field-of-view.[175]

However, the impact of these limitations can be greatly reduced, and in most cases rendered inconsequential, by giving them due consideration when designing content for the AVIE. The AVIE, like any medium, possesses idiosyncrasies, and these must be accommodated in the creation of content. Over the years, a body of knowledge of how to effectively design for the AVIE, to overcome and even exploit its limitations, has been amassed within iCinema; a body of knowledge encoded in the numerous works of art that have been successfully exhibited on the AVIE platform, including *La Dispersion Du Fils.*

3.14.1 Improvements and Future Work

There are number of avenues for improving the AVIE. Resolution can be readily improved with new projectors, and as 4K projectors rapidly become standard, an overall resolution of 16000x2000 pixels could be attained with just 5 projectors. Even

---

[175] And in the case of the older passive-stereo AVIE configurations, poor stereo-contrast at the bottom of the screen.

more interesting would be sixteen 4K projectors, each turned on its side, giving a total resolution of 32000x4000 pixels.

The solution to the problems of image brightness, stereo separation and raised black levels, may lie in one promising development: laser projectors. In particular, if laser lamps can be used to implement spectral comb-filter stereo without wastage of light, then a significant increase in stereo image quality can be anticipated (Jorke & Fritz, 2003). It is however, unclear how the problem of inter-reflected light might be resolved with a projection-based system. Here, perhaps emissive screen technologies (e.g. LED, OLED or even back-projection) fitted with anti-reflective one-way barriers may provide the answer.[176]

The problem of glare, where light is projected into the eyes of viewers when they stray into the "shadow-free zone", can be solved by using data obtained from the tracking system to project black onto the eyes (and if desired, the whole body) of the viewers.

For the limited vertical field of view, it has already been observed that by simply increasing the height of the screen with respect to the radius, a greater sense of immersion is attained. If the height can be extended sufficiently, then the audience could even be placed on a raised platform, providing a field of view that extended below the ground plane, as was very often the case with the AVIE's 18th and 19th century ancestors. Beyond this, however, a greater sense of immersion may only be possible by providing each viewer with truly correct imagery for their point of view. To achieve this, advances in image multiplexing, be it by time or wavelength, will be required (Fröhlich et al., 2005).

For the audio component, wave-field synthesis presents one possible path to greater spatial fidelity.

---

[176] But only once a truly seamless image is possible. See Febretti et al. (2013) for a very recent implementation of cylindrical display using LCD panels. That the inventors of the original CAVE system have abandoned the cubic screen for a cylindrical format in the creation of the 'CAVE2' can be considered further proof of the viability of the AVIE's design.

Recently, it has become possible to equip a single workstation with 4 graphics cards, allowing a single machine to drive 16 or even 24 displays. This suggests an avenue for reducing the complexity of AVIE software by removing the need for clustering altogether. It should be noted, however, that such a single-computer system would still require optimal use of multiple multi-core processors and GPUs, which from the point of view of software complexity, is not significantly different to using multiple computers. Most importantly, if the 'slice-rendering' technique for rendering omnistereo is used, then a method for rendering all slices simultaneously must be ensured. As for the omnistereo algorithm itself, improvements include dynamic auto-adaptation of slice-width, vergence and binocular disparity over the surface of the cylinder and over time. Another welcome addition to the system would be an automatic projection calibration solution, where feedback from a camera is used to automatically calculate the distortion correction and blend parameters.

On the software side, the AVIE faces a new challenge. *Virtools* is now a discontinued platform, having outlived its lifespan and the affections of its maker. A move towards a new software platform is underway, but it is surprising to note that, as of 2013, there are no high quality game or simulation platforms, commercial or otherwise, that natively support clustering. This wheel, it seems, must be invented once again.

3.14.2 Conclusion

Having met the requirements set out in Section 3.4, the AVIE can be considered a viable and effective platform for the presentation of multi-user immersive experiences. With twelve AVIE systems in operation around the world, it has a proven robustness and reliability. It has been exhibited publically in exhibitions and festivals worldwide, proving both its durability and portability. Indeed, the entire system can be unpacked, installed and calibrated for use in no more than three days. It has a proven versatility, having hosted a wide variety of immersive applications and, with the iCinema SDK providing a powerful and proven development platform, permits rapid development of new content. It is not without flaws, but these flaws can be readily accommodated in the design of immersive applications. It is perhaps unique among virtual reality platforms for its capacity to immerse up to 20 viewers in a shared physical space, untethered and free to move at will, and free to interact not only with the virtual world, but with one another. In this respect, the AVIE provides a platform for shared immersive mediated experiences that could not be delivered on any other platform.

4. *La Dispersion Du Fils*

4.1 Introduction

This section concerns the third major component of this thesis, the artwork entitled *La Dispersion Du Fils.* Designed specifically for display in the AVIE, *La Dispersion Du Fils* serves as a vehicle for the exploration of the aesthetic and immersive possibilities and limitations of the AVIE platform and, by extension, of panoramic cinema in general. The ideas and observations put forward in the first part of this thesis arise from, and are explored through, the process of creating this artwork. Particular attention is given to the roles of ego-motion and vection, interactivity, realism and plausibility, audio-visual binding and perceptual rest-frames and the role they play in the invocation of presence.

The work can be seen as an ongoing series of experiments, some successful, others less so. Throughout the development of the work, which began in June 2008 and continues to the present day, the work has been continually reshaped and augmented. And with no end-point or final destination envisaged, the work is destined to remain permanently in flux. To date, the work has been exhibited publicly on nine different occasions,[177] with each exhibition representing a snapshot of the work as it stood at that time.

La Dispersion Du Fils is the result of a collaboration with Jean Michel Bruyère, Delphine Varas and Thierry Arredondo, members of the Marseille-based artist collective LFKs. It is a continuation and culmination, indeed the climax, of the *VøSPAZÀR cycle*, a decade-long inquiry into the classical myth of Actaeon and Diana.

The work takes the form of a journey, or descent, through structures and landscapes constructed wholly from moving images. These images are drawn from the vast library of filmic material created by LFKs during the course of the VøSPAZÀR cycle – over 500 short films created between 1999 and 2007. *La Dispersion Du Fils,* in uniting and presenting these films as a whole, transforms them into a form of living "cynematic"

[177] See Appendix A for details of these exhibitions.

library,[178] a vast and boundless zoetrope within which the viewer is variably engulfed, encircled, or entombed.

The films of LFKs therefore play a central role in the work, for they form the basic building material from which the virtual world is constructed. These images are taken as indivisible atomic elements from which larger structures are composed, but rather than arranging images temporally (as is done in the traditional art of film editing), the images are conjoined spatially to form three-dimensional surfaces. Everything visible in *La Dispersion Du Fils* is constructed from the images contained within this library of films.

La Dispersion Du Fils has been created in accordance with the methods and principles adopted by LFKs in the creation of all their works. As a consequence, great emphasis has been placed on the roles of chance and discovery. To this end, the work is real-time, generative and *stochastic*, in that the system evolves in a largely non-deterministic and unpredictable manner. It never repeats and never ends; every moment is, in some respect, unique and destined never to be visited again. Indeed, of the innumerable potential states the system might adopt, only a vanishingly small fraction of these states will ever be seen. Both the metamorphosis of Actaeon and the subsequent fate of his hounds are captured in this restless and unending meandering through the space of all possible states.

4.2 BACKGROUND

In order to understand how *La Dispersion Du Fils* came to take one particular form over another, a certain historical perspective is required. There are some aspects of the work, both technical and aesthetic, that may only be understood by knowing the story of the work, its precursors, and how the collaboration came to be, all of which I will endeavour to outline here.

[178] The term is a play on the Ancient Greek *kynikos* (κυνικός), from which the English word 'cynic' derives. This reflects not only the dog-like nature of the work, but the strong connection with the philosophy of the Cynics, for whom the dog was an icon. Bruyère's works and writings are full of such plays on language.

In November of 2007, Jean Michel Bruyère and Delphine Varas, members of the Marseille-based artist collective LFKs, visited the iCinema Centre for the first time. The purpose of their visit was to experience the AVIE system and assess the possibility of Bruyère and LFKs creating a work of art for display in the AVIE system. Their visit had been organised by Richard Castelli, founder and director of Epidemic,[179] curator and producer, who counts among the artists he represents Ulf Langheinrich and Kurt Hentschlaeger, (and their former collective identity, Granular Synthesis), Édouard Lock, Saburo Teshigawara, Robert Lepage, Thomas McIntosh and Emmanuel Madan, Dumb Type, as well as Bruyère and the LFKs, and, here is the link, Jeffrey Shaw, who in 2007 was the director of the iCinema Centre. Castelli, having previously seen the AVIE system himself, had described it to Bruyère who subsequently expressed a desire to create an artwork with the system.

This Castelli-Shaw-Bruyère constellation was not without precedence. Shaw's opus is largely defined by the invention of novel immersive platforms, and many of Shaw's works can be easily and precisely divided into *platform* and *content*. The platforms tend to share a singular quality: they are all easily re-purposed. They are neither suggestive of, nor implicitly restricted to, the depiction of any particular themes, stories, events, styles or genres.[180] This property of Shaw's work is evident in the *Legible City*, *The Golden Calf*, the *PLACE* series and *Eavesdrop*, *Unmakeable Love*, and continues through to the developments made at iCinema: *Conversations*, *Conversations at the Studio*, *Brother Where Art Though* and the AVIE-based projects such as *TVisionarium*. Often Shaw's platforms are given their own names – EVE-dome, iDome, SphereCam, VROOM/Reactor – and go on to enjoy prosperous lives of their own, beyond the scope of any one artwork or artist.

However, while Shaw's creative genius may lie in the creation of novel platforms and interfaces, his works cannot be described as "interface art" – the form of interactive media art pioneered by Myron Krueger (1977; 1985) and David Rokeby (1995, 1998) and

[179] See *epidemic.net*

[180] In some cases it is Shaw himself re-purposing an existing platform. For example, *Configuring the CAVE* (1996) and its later reincarnation *Reconfiguring the CAVE* (2001) are both built upon the CAVE virtual reality platform.

now continued *en-masse* by the *Kinect* generation. “Interface art” is conspicuous for its distinct lack of “content”; or to quote Rokeby, “interfaces are content” (Rokeby, 1998, p. 1). By contrast, Shaw’s interfaces serve only as devices for entering into a world, narrative or idea. Shaw’s platforms need “content” in exactly the same way a cinema needs a movie:

> Here as well, the cinema is both the model and inspiration. In the cinema, a technological apparatus was invented: film, the camera, the projector, etc, which innumerable artists have used to craft completely personal statements. Many of the ‘machines’ I have developed also have this almost generic capability to become expressive tools in other artists’ hands. *AVIE*, for instance, is a paradigmatic contemporary environment for the expression of panoramic spaces of representation, which follows in the traditions of the Baroque’s immersive surround and of panorama painting. As an artist, I see myself both as a creator of new systems of representation, which are made even more significant and valuable by other artists’ use of them, as well as the creator of unique instances of representation using these systems.
>
> *Jeffrey Shaw, in interview with Laurent Catala (Shaw & Catala, 2013)*

On two previous occasions Bruyère and LFKs had created works using a platform created by Jeffrey Shaw. First, in 1999, Bruyère was commissioned to create a film for the Extended Virtual Environment, or EVE Dome, which had been created by Shaw many years earlier, in 1993, during his directorship of the ZKM Institute of Visual Media in Karlsruhe, Germany. The outcome of this commission was the work *Si Poteris Narrare, Licet*, inaugurated at the *Festival Via* in Maubeuge, France, in March 2002. The second instance occurred in 2007, when Bruyère was invited by the ZKM Institute of Visual Media, this time to create a work using their panoramic camera and projection systems initiated under Shaw’s directorship and further developed by Bernd Lintermann. The result of this commission was the work *$CaMg(CO_3)_2$*.

And so, with all the elements falling into place - most notably the support and patronage of Castelli and an invitation from Shaw and iCinema - Bruyère and LFKs would be invited to imagine, for a third time, a new work for an existing platform.

4.3 JEAN MICHEL BRUYÈRE AND LFKS

To describe the works of Bruyère and LFKs is a challenging task. The first obstacle an aspiring biographer will encounter is a scarcity of writing on Jean Michel Bruyère and LFKs and their works, and an equal paucity of publicly available documentation of their work. This is somewhat at odds with the volume of creative output generated by Bruyère and company over the years, which includes almost continuous exhibitions and performances in the major art and theatre festivals and venues of Europe and Asia. To what degree this is due to Bruyère's reticence for all forms of publicity and promotion, or to his refusal to speak publicly about himself or his work, or more to his political views - an open rebelliousness which often manifests in scathing critiques of institutional thought and structure, including art criticism, academia and the press - is hard to determine. Or perhaps it is more attributable to the sense of cynicism, irreverence and humour with which he approaches his role as 'artist' in the 'art-world.'[181] Most likely it is for all these reasons that very little writing about his work can be found through the ordinary channels. In a world where many artists could be described as 'global brands,' and where the publicity of an artwork is often more involved than the artwork itself, it is rare and refreshing to encounter one as prolific, apparently successful and at the same time, anonymous.

With a career spanning 35 years, a thorough review of the creations of Bruyère and the LFKs would be a rich, rewarding endeavour. His travails include such things as the foundation of the *Man-Keneen-Ki* in 1996, a home – medical clinic – contemporary art school dedicated to the street children of Dakar,[182] the publication the art magazine *NK 1314* (1994 - 1999)[183] and over a dozen books and texts, art direction and set design for such institutions as the *Institut du Monde Arabe, Musée Dapper, Ballet Atlantique-Régine*

[181] "We are under no obligation to even pretend that we are able to define what we do, or will do." Jean Michel Bruyère in response to Antoine de Baecque. *(« nous n'avons pas d'obligation à nous prétendre capables de définir ce que l'on fait, ce que l'on va faire ou fera »).* Festival d'Avignon 2009 program notes.

[182] Man-Keneen-Ki maison / clinique / école art-contemporain des enfants et jeunes errants de Dakar. *www.sklunk.net/spip.php?article241*.

[183] Within which a collection of the writings of Bruyère can be found. See *www.sklunk.net.*

Chopinot and *Ballet d'Europe*, and not least, the founding in 1986 of the LFKs. Alas, such a review is not something that can be undertaken here. Instead, I offer just the briefest of introductions, with a focus on the elements relevant to this thesis.

The oeuvre of Jean Michel Bruyère and LFKs spans prose and poetry, photography, painting and graphic design, sculpture, music and film, concerts, installations, performance, mechatronics, set design, theatre and even opera. Officially based in Marseille, but with members spread from Senegal to Poland, their works are presented at major festivals world-wide. The group (originally known as *La Fabriks* but today more often referred to as LFKs), comprises a musician, composer, singer, actor, seamstress and costume designer, film editor, photographer, carpenter and builder, philosopher, cook and sometimes a doctor.

With the exception of a handful of works (*Si Poteris Narrare, Licet* and *CaMg[CO*$_3$*]*$_2$ and *La Dispersion Du Fils*), the works of LFKs could not be described as media art, and computers are only occasionally, and never overtly, employed in their work. Nonetheless, themes that are often central to computer-mediated art, and are certainly central to this thesis, feature prominently in their work. These are immersion, interaction, stochasm (the role of chance and indeterminism) and simulation, not in the sense imitation, but in the procedural/processual sense of the word.

A notion of immersion pervades the work of LFKs. First, it describes, in a metaphoric sense, the very creative process from which the works surface. A theme is extensively and exhaustively researched over a period of years, during which countless physical objects and artefacts are amassed, libraries of reference material collected and catalogued, and threads of inquiry followed without direction or destination. Slowly, (and only once a sufficient density or entropy of information is achieved, and all obvious avenues exhausted and abandoned), a world forms, replete with its own characters and narratives, language and literature, mythology and philosophy.

From this ongoing process, the public works of LFKs emerge, like islands rising from the seabed - the visible portions of a sunken landscape. These works - which can loosely be described as performance/installations - often involve the transformation of a large building or venue into a self-contained microcosm, populated with images, text, sculpture and sound installation, cinema, music, and performers. The audience is invited to wander freely through this world, at their own pace and leisure. The performances have no fixed beginning or end, often running eight or ten hours a day, for a number of days or weeks in succession.

Figure 68 - Le Préau d'un Seul. Bruyère & LFKs. Festival d'Avignon 2009. Issa Samb is the sole detainee in an immigrant detention centre. During the performance, which runs 10 hours per day for 10 days,the audience are free to wander at will throughout the camp.

Although the work is theatrical, it is rarely scripted. The members of LFKs are never asked to adopt fictitious persona, to mimic or pretend to perform an action. Rather, the actors are given tasks and functions, and they simply go about fulfilling these responsibilities. For example, in *Le Préau d'un Seul,* a sprawling and intricately detailed immigrant detention centre is constructed to house a single solitary detainee - the Senegalese poet philosopher Issa Samb.[184] Within this camp, for ten hours a day over ten days, a real cook cooks, a real doctor performs medical examinations, while

[184] "In 1974 Samb founded, together with a group of artists, writers, filmmakers, performance artists, and musicians, the Dakar-based *Laboratoire Agit'Art*. The aim of the group was to transform the nature of artistic practice from a formalist, object-bound way of working to practices that were based on experimentation and agitation, on process rather than product, ephemerality rather than permanence, and political and social ideas over aesthetic ones. Focusing on communication between the artist and audience over physical objects, the actions of *Laboratoire Agit'Art* engaged with the immediate sociopolitical situation" (Njami et al., 2013). A long time companion and collaborator of Bruyère's, it is clear that Issa Samb has influenced Bruyère's approach to life and art.

seamstress sews new uniforms for the detainee. A team of press officers produce daily gazettes. Throughout all this the prisoner is ordered about, weighed and measured, fed his rations, dressed, undressed, all before the passing audience. But not all activities are so easily understood, for in addition to this element of realism, a surrealist and often absurdist aesthetic prevails. That is, even as some actors set about their tasks repeatedly and conscientiously, the purpose and meaning of these actions often remain unfathomable. Bruyère has previously referred to this approach as "théâtre documentaire" (Gerz, 2000); an approach to theatre that shares much in common with the "ethno-fiction" or "documentary surrealism" of film-maker and ethnographer Jean Rouch (Rouch & Feld, 2003).

Figure 69 - Le Préau d'un Seul. Bruyère & LFKs. Festival/Tokyo 2012.

With instructions but no script, a world is created, and then left to run its own course. If the terminology of media art is adopted, the works could be described as generative or procedural. A world is populated with agents and artefacts, and forces, logic and constraints, which together define a form of evolutionary dynamics - a *physics*, so to speak – which propels the agents on evolutionary trajectories.

Most importantly, this evolution is non-deterministic. The performances of LFKs never repeat and seldom reach any form of denouement. There is no pre-conceived destination towards which the world evolves. Chance and discovery play a pivotal role in the methods of LFKs, a stochastic character to the work that is often evident in the kinetic sculptures and contraptions, sound works and video projections that often populate their installations.

Figure 70 - Le Préau d'un Seul - Autonomous painting machine. Bruyère & LFKs. Festival d'Avignon 2009. The kinetic contraption paints with sponge and pig's blood.

For example, in the *Le Préau d'un Seul* performed in Avignon 2009, a lurching, mechanical arm paints with pig's blood and a sponge, blind to its own awkward movements. A row of animated medical beds perform an algorithmic ballet, composing a mechanical score with their squeaks, whirrs and pneumatic sighs. In a more recent project, *Bobby Seale Got His 9 - VitaNONnova #3* (2012), a colossal projection screen shows a never-ending film. The film is dynamically and algorithmically editing itself, drawing clips from a database of 24 hours of footage according to some non-deterministic reasoning.

Through all of this, the audience are free to wander throughout the space at will. As argued in the beginning of this thesis, the capacity to navigate freely through one's environment is critical to any sense of immersion. In this fundamental sense, the works can be said to be interactive, as the viewer is left to his or her own devices to discover the world, and construct their own paths and narratives.

Figure 71 - Le Préau d'un Seul - The Path of Damastes. Bruyère & LFKs. Festival d'Avignon 2009. The beds move in a synchronised, algorithmic dance.

This interactivity, however, is almost always limited to navigation. This is not a rule, but a behaviour that seems to emerge naturally, for the viewers receive no instruction and are free to behave as they wish. However, they are never given nor asked to adopt fictional roles, and while the performers may react to the presence of the audience, they seldom engage the audience directly. The result is that the audience do not tend to have, *nor seek to have*, direct influence on the unfolding narrative, and the experience is seldom an exercise in 'participatory theatre.' For the audience, just as it is for the members of LFKs themselves, their role in the narrative is defined entirely by their behaviour and actions, and not the other way around. They are to the narrative exactly that which their actions dictate - spectators, voyeurs, witnesses, trespassers, complicit bystanders.

This freedom to physically explore demands of the fictional world a certain completeness, concreteness, consistency and self-containment, if presence is to be maintained. The viewer must never be confronted with events or boundaries that appear contrived or artificial, or inconsistent with the world's own inner logic.

These are just some of the qualities of the works of LFKs, and by no means the most important (for no mention is made of the political, poetic or philosophical nature of their work). However, in drawing attention to them, I hope to illustrate that the mechanics of presence, with its notions of immersion, ego-motion and interaction, processualism and simulation, plausibility and realism, easily extend beyond the domain of computer-mediated art.

4.4 THE TRAGEDY OF ACTÆON

> Nunc tibi me posito visam velamine narres, si poteris narrare, licet
>
> Ovid, *Metamorphoses. III. 192-193*

So said Diana to Actaeon, the young hunter who stumbled by error upon her bathing and dared contemplate her naked beauty. *Si poteris narrare, licet*. If you can tell, do so; I give my consent. With these words however, Diana gives not license, but passes sentence. With a splash of her bathwater the Goddess transforms the hapless Actaeon into a deer; her feigned consent a final taunt before she abandons Actaeon to his fifty dogs.

> This said,
> She to his neck and ears new length imparts;
> T' his brow the antlers of long-living Harts
> His legs and feet with arms and hands supply'd;
> And cloth'd his body in a spotted hide.
> To this, feare added.
>
> *Metamorphoses. III. (Ovid & Sandys, 1632)*

With an unfamiliar fleetness of foot, Actaeon takes flight, and for some time keeps ahead of his hounds. But he cannot escape, for they were trained in the art of the hunt by Actaeon himself, and know the mountains better than even he.

> The whole pack, with the lust of blood upon them,
> Come baying over cliffs and crags and ledges
> Where no trail runs: Actaeon, once pursuer
> Over this very ground, is now pursued,
> Fleeing his old companions. He would cry
> "I am Actaeon: recognize your master!"
> But the words fail, and nobody could hear him
> So full the air of baying.

Metamorphoses. III. (Ovid & Humphries, 1955)

Actaeon, whose transmutation has left his mind untouched, calls out to his faithful hounds. *Melampus! Hylaeus! Labros! Aëllo! It is I, Actaeon!* He need only say a word, any word, to make himself known, but all he can evict is a mournful, guttural bray, not the voice of a man, but not altogether the bellow of an animal. Too late! His hounds are upon him.

> And all together nip and slash and fasten
> Till there is no more room for wounds. He groans,
> Making a sound not human, but a sound
> No stag could utter either, and the ridges
> Are filled with that heart-breaking kind of moaning.
> Actaeon goes to his knees, like a man praying,
> Faces them all in silence, with his eyes
> In mute appeal, having no arms to plead with,
> To stretch to them for mercy. His companions,
> The other hunting lads, urge on the pack
> With shouts as they did always, and not knowing
> What has become of him, they call Actaeon!
> Actaeon! each one louder than the others,
> As if they thought him miles away. He answers,
> Hearing his name, by turning his head toward them,
> And hears them growl and grumble at his absence,
> Calling him lazy, missing the good show
> Of quarry brought to bay. Absence, for certain,
> He would prefer, but he is there; and surely
> He would rather see and hear the dogs than feel them.
> They circle him, dash in, and nip, and mangle

And lacerate and tear their prey, not master,
No master whom they know, only a deer.

Metamorphoses. III. (Ovid & Humphries, 1955)

The story of Actaeon, this unfortunate hunter transformed into a deer by a spiteful Artemis (Diana) and torn apart by his own dogs, has seduced two millennia of artist, writer, painter and poet. Its themes of metamorphosis and transcendence, forbidden knowledge, truth and the pursuit of it, yearning and sacrifice, transgression, injustice and wrath, chance, fortune and error have inspired the imaginations of Ovid, Petrarch, Dante, Gower, Shakespeare, Jung, Lacan, Sartre and Klossowski, as well as the painters Titian, Rembrandt, van Dyck, Turner, Delacroix, Cesari, Cranach (the Elder and the Younger), and the poets Wharton, Brown, Pound, Hughes and Heaney.

Figure 72 - Diana und Aktäon, Cranach the Younger, c.1550, Hessisches Landesmuseum, Darmstadt

Among the countless interpretations of the myth, none are so pertinent or potent as that offered by Giordano Bruno, for whom the myth expressed the philosophical quest for knowledge, truth and beauty.

> Thus Actaeon, with these thoughts, these dogs that sought goodness, wisdom, beauty, the sylvan wild beast outside themselves, and in that way he caught up with it, enraptured by such great beauty, he himself became the prey, seeing himself turn into what he was chasing; and he noticed that in the eyes of his dogs, in his thoughts, he became the coveted prey, for, having already assumed divinity within himself, there was no need to look for it on the outside.
>
> Giordano Bruno, *De Gli Eroici Furori* (1585), translated in Ordine (1996, p. 88)

Of the sacrifice entailed in the pursuit of truth, no one can speak with greater authority than Bruno. Martyred for his convictions, Bruno's reading of the myth is terribly prescient.

> And so we see Actæon, pursued by his own dogs, persecuted by his own thoughts, running and forging a new path. His strength is renewed to proceed divinely and with lighter steps... into denser thickets, into deserts, into the region of incomprehensibility. [...] Here his many big dogs put him to death; here ends his life in the eyes of the mad and sensual world of blindness and illusion, and he begins to live on the plane of the intellect; he lives the life of gods, feeds on ambrosia and becomes drunk on nectar.
>
> Bruno (1585), translated in Ordine (1996, p. 88)

4.4.1 The VøSPAZÀR cycle and Art Cynique

Beginning in 1994, Bruyère and the LFKs adopted the tale of Diana and Actaeon as a form of manifesto; an approach to the practice of art that would influence their work for decades to follow. The misfortune of Actaeon would be taken as the founding act of *art cynique*, named not so much for the cynicism expressed by Diana when she condemned Actaeon to his death, but in memory of the fifty dogs of Actaeon. The artists would embark on a "permanent and unending reflection" on the perdition of Actaeon and wrath of Artemis. Above all, however, they would concern themselves with the fate of the dogs, hitherto ignored and forgotten by history. Pall bearers of Actaeon's body (for they carried him in their bellies) and sole witnesses to his demise (it can only be through them that we know of Actaeon's fate today), they are all that remained of the hunter. What became of them? What did they make of their new found liberty? For a long time they searched the mountains for their beloved master, but he was to be found nowhere. Yet he was with them, within them, coursing through their veins, and as they wandered throughout the forest, they unwittingly scattered him in their shit and piss. Like this, wherever they went they found traces of his being, which fuelled within them an ever increasing frenzy, for he seemed to be truly nowhere and everywhere.

In adopting *art cynique*, the artists of LFKs declared themselves "to be dogs;" distant descendants of the hounds of Actaeon.[185] Having eaten the twice-transformed flesh of their master (transformed first from man to deer, and then from deer to motion), Actaeon's dogs were themselves subject to transmutation, over generations becoming the VøSPAZÀR: a "cynomorphic people" of strays and vagabonds, wayward in language and learning and destined to forever roam aimlessly and endlessly.

In adopting *art cynique*, the artists of LFKs renounce "both coherence and narration" in their work and with it the comprehensibility of language, for its inability to say the unsayable. Instead, they adopt an "art of silence" and recognize only the "value of the pure deed, for which they seek no end other than itself" (Bruyère, 2003).

4.4.2 The Films of LFKs

Figure 73 - Si Poteris, Licet - 9 Guerriers (LFKs, 2002).

[185] See Appendix C for an example of the writings of the VøSPAZÀR.

Between 1999 to 2007, the LFKs would develop and present twenty unique projects - exhibitions, installations or performances - in the course of their dogged inquiry into the fate of Actaeon and his dogs.[186] Many of these projects would involve the creation of films, and so, by the end of this époque, the LFKs had amassed a vast library of film and video material. Diverse in style and content, almost all the films were created for display within installations and performances, where they would often be accompanied by sculpture and live performance. Many of the films were originally presented as polyptychs; multi-screen installations dispersed physically in space. A significant number of the films were created in Senegal and feature the young students of the *Man-Keneen-Ki* shelter/clinic/art-school founded by Jean Michel Bruyère in Dakar.

Figure 74 - L'Insulte Faite Au Paysage - Fioretti (LFKs, 2005).

In preparation for *La Dispersion Du Fils,* the library was carefully edited by Delphine Varas to yield a database comprising 527 unique films. 92 are black and white, with the rest in colour. 142 have an aspect ratio of 16:9, while the remainder are 4:3, with the exception of one - the panorama *$CaMg[CO_3]_2$*. In total, more than 12 gigabytes of film data, enduring 18 hours and 44 minutes were one to watch it all serially, constitute the film database of *La Dispersion Du Fils.*

---

[186] See Appendix D.

Figure 75 - L'Insulte Faite Au Paysage - Transes Issa (LFKs, 2005).

4.5 TECHNICAL HERITAGE

Being a collaboration between LFKs and iCinema, *La Dispersion Du Fils* would be not just a continuation of the work of LFKs, but also a continuation of the research and development taking place at iCinema. In this section the pre-existing projects that most influenced the initial concept of the project are discussed.

Towards the end of 2007, Bruyère and Varas were invited to spend some time getting acquainted with the possibilities and limitations of the AVIE system. Upon our first meeting, an effort was made to explain that the device was "an open system" capable of simulating almost any experience or situation. That is, while some phenomena or situations may be more accurately, more easily or more effectively portrayed in the AVIE, there is nothing that the system is inherently incapable of portraying.

A fine balance was needed, for what they would see during their visit should impart an idea of the *potential* of the AVIE, without restricting or unduly influencing their understanding of the kinds of experiences that might be possible. To understand the precariousness of this moment, one might imagine the same dilemma for an inventor of a digital synthesiser. The essential property of his or her new machine is not that it can mimic any earthly sound, but that it can synthesise new, hitherto unheard sounds. The inventor invites a composer to imagine new works for his new machine, but in order to demonstrate it, can only play a tiny fraction of all the possible sounds the machine might issue. These examples inescapably influence the composer's ideas of what the machine sounds *like*, the very idea of which is in complete opposition to the essence of his invention. Facing this, the inventor has two possible remedies: demonstrate a greater diversity of examples, or provide the composer with an understanding of how the machine functions, both in the hope that the composer arrives at an understanding that his machine does not, in fact, sound *like* anything at all. At the time of this first visit, only a handful of working AVIE applications existed: *TVisionarium,* the Mining VR simulator, and the Panorama Player. In retrospect, it is clear that all of these had some influence on the imagination of Bruyère and Varas and the initial design of *La Dispersion Du Fils*, in a number of ways.

From the Mining VR prototype, two essential lessons were learned. The first was that the most basic and common method for synthesising real-time images for a regular screen - the lighting and rasterisation of textured polygons - was a viable method for creating imagery for the AVIE, and that, even when using such basic techniques, virtual three dimensional space and structure could feel tangibly solid and present when presented in the AVIE. This was important, for it meant that the success of the AVIE did not depend on the use or invention of new or exotic rendering techniques.

The Mining VR prototype also suggested the immersive potential of *simulated ego-motion.*[187] The flight of the virtual camera through the long corridors and mine shafts of the simulated mine, aroused in the audience sensations of physical movement;

[187] The link between vection and presence had been hinted at previously in one of the earliest AVIE experiments, the Panorama Player. The first images shown in the AVIE were omnistereoscopic photographs created by photographer Peter Murphy, a specialist in panoramic photography. Murphy had captured his panoramas using the slit-mosaic method with a single rotated off-axis camera. Now, as it happened, some of Murphy's panoramas had been captured using a very wide fish eye lens, and the resulting panoramas were essentially spheres. As the AVIE has a comparatively narrow vertical field of view, this permitted a vertical translation of the camera, a feature implemented by the author when he created the *Panorama Player*. At the onset of the levitation, the audience would almost always issue a collective gasp, for the introduction of this very simple and subtle movement had an unexpectedly profound effect: the movement of the image was immediately and physically perceived as movement of oneself.

it was not a camera but themselves that were moving. Interestingly, this was not the case with the panoramic video, for despite depicting a drive through the streets of Sydney, the sense of vection was never as powerful as that in the Mining VR simulator. The reason can be attributed to the panoramic video being monoscopic, rather than stereoscopic, and also to the frame-rate of the images, for the panoramic video was limited to 25 frames per second, while the Mining VR simulation typically run at 60.

From *TVisionarium*, two insights were gleaned. The first was that it is possible technically, as well as ergonomically, aesthetically and conceptually, to display traditional two-dimensional film material in the AVIE. This discovery, when cast in the light of the large repertoire of film material created by LFKs, can perhaps be seen as the initial spark for the project. *TVisionarium* demonstrated an approach to the immersive display of a large body of 2D film material, and it is quite reasonable to assume that a simple re-purposing of the *TVisionarium* system into some kind of virtual and interactive LFKs film archive was the expected result of the initial commission.

However, it was a different feature of *TVisionarium* that may have had more influence on the initial design of *La Dispersion Du Fils*. *TVisionarium* demonstrated that by distributing the images in three dimensional space, it is possible to construct a sense of space, or even *sense of place*, inside the AVIE, by treating film as physical object. It is in this discovery and the suggestion of heightened immersion through vection that the seeds of *La Dispersion Du Fils* can be found.

And so, on the final day of this initial visit in 2007, a rapid experiment was made. A random collection of LFKs films, about twenty in number, were distributed across the surface of polygonal torus. The torus had a tubular diameter of around 8 metres, sufficiently large to pilot the AVIE throughout. The result was the sensation of travelling through an endless tunnel - a tunnel constructed from images - an embryonic version of what would become known as the *Helix*.

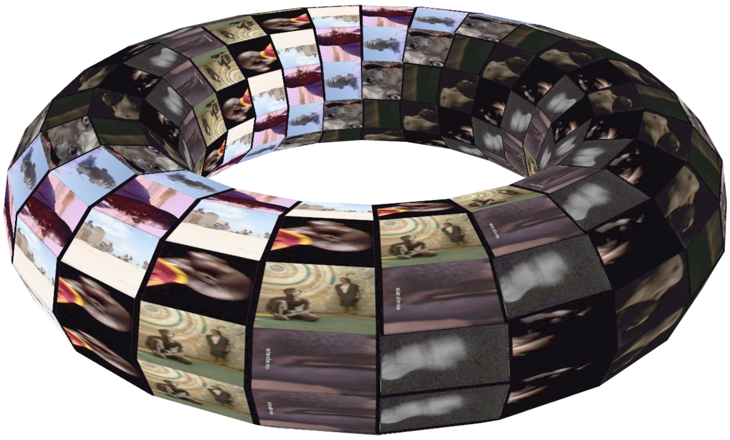


Figure 76 - Video Torus, precursor to the Helix.

4.6 The Helix

The central element of *La Dispersion Du Fils* is the Helix. An endless, undulating sinuous tunnel constructed entirely from moving images, the Helix exists in a state of eternal metamorphic flux.

The essential experience of *La Dispersion Du Fils* is a journey through the Helix. The viewer experiences the Helix by travelling through it, within it. In this manner, the Helix serves a conduit, or path along which the traveller is propelled.

The Helix is constructed entirely from moving images. The organisation of these images, their size, orientation, arrangement and intermittency are in constant evolution, presenting the traveller with an ever-changing matrix of images as they progress through their host.

The roles played by the Helix are manifold. As a structure, it serves as path and conduit, along which the viewer is propelled. As a mosaic of images, it serves as a medium; a form of living mosaic of signs and symbols, memories and prophecies. (To read the images in the skin of the Helix, then, is nothing other than the practice of haruspicy – the search for signs or omens in the entrails of animal.) In a perpetual state of growth and decay, it advances along serpentine and meandering paths, delivering the viewer to and from the various 'extra-Helicular' phenomena that populate *La Dispersion Du Fils*. In this respect, it serves as vessel and vehicle, companion and host.

And as an animate being, possessed with both mood and mind of its own, the Helix serves somehow as both protagonist and antagonist in the unfolding narrative.

Figure 77 - The Helix. Photo taken in AVIE.

Presented here is an account of the construction of the Helix, and the manner in which it employs vection, animation and spatial audio to create a highly immersive experience.

The Helix is composed of two fundamental parts: the *spine* and the *skin*. The spine determines the overall shape, twist and tumble of the Helix, and defines its trajectory through space. It is invisible. The skin is the visible manifestation of the Helix. It is the surface spun around the spine and rendered visible by light.

4.6.1 The Spine of the Helix

Running through the centre of the Helix is a sequence of bones.[188] Each bone is defined by a position, orientation and length. Under Catmull-Rom spline interpolation (Catmull & Rom, 1974), the bones are transformed into a smooth 3D curve with C^1

[188] 500 bones, each 4 metres in length. Early versions (2008-2010) used only 250 bones, due to the limited computing power of the computers at the time. It is expected that the next iteration of the work will increase the bone count to 1000.

continuity. This curve, 2 kilometres in length, is the spine of the helix, and it is the foundation of the Helix system.

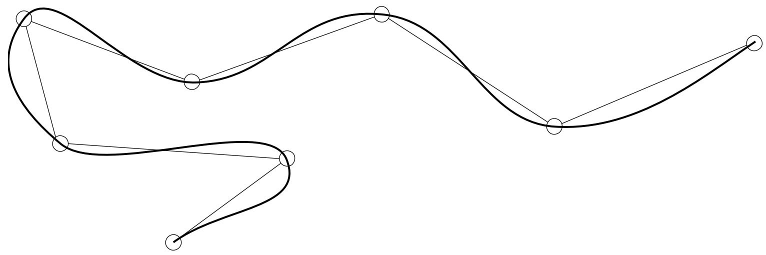


Figure 78 - Catmull-Rom spline interpolation.
Given any a set of points, a smooth curve is produced passing through all.

The spine is used as the basis for *helix-space,* a reference frame that proves useful for many subsequent calculations. Helix-space is 3-dimensional, with z being the position along the spine, r being the radial distance from the spine and θ the angle of rotation around the spine. Note that, the mapping from helix-space to regular Cartesian space is not one-to-one (bijective); for while a point in helix-space always maps to a unique point in 3D space, different points in helix-space may map to the same point in 3D space.

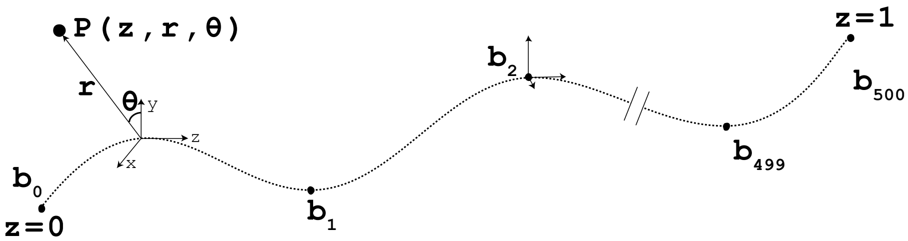


Figure 79 - Helix Space.
A point in space is defined relative to the spine by 3 coordinates: z, r and θ

Catmull-Rom interpolation also provides a tangent vector for all points on the spine which, together with the orientation of the bones, is used to calculate normal and binormal vectors. This is done using the concept of a *rotation minimizing frame.*[189] Taken together, the tangent, normal and binormal form an orthonormal basis (called here the

---

[189] Bishop (1975) first proposed the rotation minimizing frame (RMF) as an alternative to the Frenet frame, which suffers from a number of deficiencies when applied to computer graphics. Since then a number of different algorithms for calculating it have been proposed. The method used here is akin to the double-reflection algorithm proposed by Wang et al. (2008). See Wang et al. (2008) for a discussion of RMFs and comparison of different approaches.

spine-frame), and provide a definition for 'up,' 'forward,' 'sideways' for all points on the spine. The rate of change of this spine-frame provides measures of curvature and torsion.

In addition to position and orientation, all bones possess a radius. The radius for points on the spine between bones is also calculated using Catmull-Rom interpolation, where the four control-points are the radii of the 4 closest bones.[190] However, because a Catmull-Rom curve must pass through its control points, it is possible that the interpolated radii fall below some desired minimum, or even below zero. To avoid this, radii are smoothly clipped to some desired minimum (r_{min}), using a smooth quadratic falloff over a defined range.

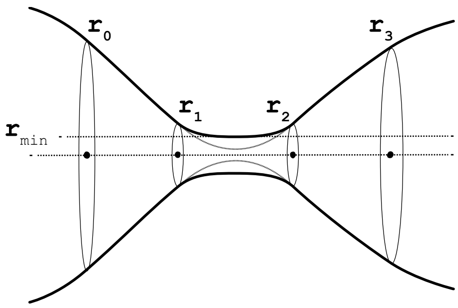


Figure 80 - Radius interpolation and smooth-clamping.
Catmull-Rom interpolation is performed on the radius, but smoothly clamped to avoid breaching a minimum radius r_{min}.

4.6.2 The Skin of the Helix

When animating the shape of an object composed of polygonal meshes, the position for every vertex must be calculated anew, each and every step in time. To reduce the complexity of this task - both in terms of computational efficiency and ease of control - the mesh can be constrained to algorithmically follow an underlying skeletal substructure. This arrangement yields a number of immediate benefits. First, an animator or algorithm responsible for setting the mesh in motion need only animate the skeleton, and need not be concerned with the vertices or topology of the mesh itself. As the skeleton is typically orders of magnitude simpler than the mesh itself, it serves as a

[190] It is possible to interpolate the 3D position and 1-D radius together in a single 4D vector (x, y, z, r), which on SIMD processor (e.g. SSE2) or a graphics card, can reduce computation time significantly.

simple and intuitive mechanism of manipulating the skin. Second, to a certain degree it mimics the anatomy of a real vertebrate, and so facilitates the mimicry of such organisms. Third, once the skeletal pose has been determined, calculation of the vertices' positions can be readily and extremely rapidly executed in parallel on a GPU, allowing for real-time animation of polygonal meshes composed of millions of vertices. For these reasons, the animation and deformation of a complex surface using a simpler 'skeleton' is a widely adopted technique in computer animation, and can be traced back at least as far as 1976 to Burtnyk and Wein's work on 2D computer animation (Burtnyk & Wein, 1976). For overviews of the technique, the reader is directed to Jacka et al. (2007), James and Twigg (2005) and Lewis et al. (2000).

Arranged around and along the spine are 32 tubular sheaths, each 62.5 metres in length. These sheaths are cylindrical polygonal meshes, each composed of 2000 rectangles ("quads"). Together they form the skin of the Helix.

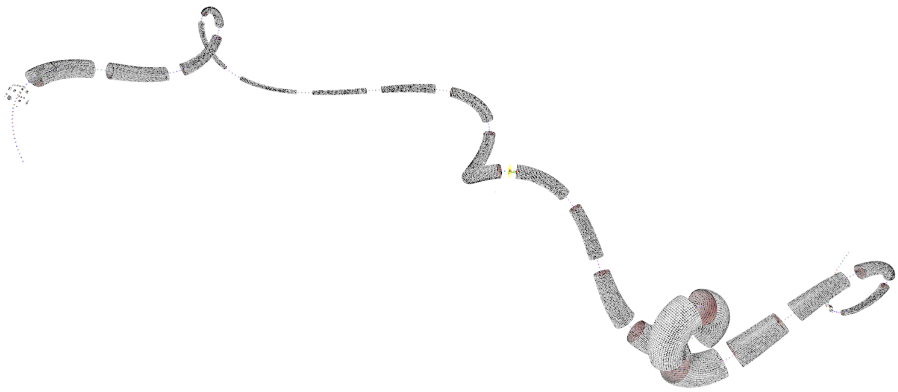


Figure 81 - Helix sections arranged along the spine.
Sections are recycled from the tail to the head as the viewer advances along the Helix.

The skin-sections can be arbitrarily positioned anywhere along the spine, with or without gaps between them. The reason for using multiple sections rather than a single long cylindrical mesh is twofold. First, it permits high-level culling of the mesh. Very often only a fraction of the helix is visible on any one computer, so having the skin divided into sections allows efficient culling of whole mesh sections on the CPU. Second, it is a necessary part of the mechanism used to create the illusion of an infinitely long tube. As the viewer travels along the helix, both bones and skin-sections are removed from the rear of the helix and replaced at the head. This recycling of bones and sections removes the need for repeated creation and destruction of resources (GPU vertex and texture buffers), which can dramatically decrease smoothness of motion and lead to fragmented memory.

This cylindrical skin is now bound to follow both the curvature and radius of the spine. Every time-step new bone-states are calculated on the CPU and then passed to the GPU as a floating point texture, where they are then used to calculate the deformed positions of the skin vertices. The difference with regular skeletal/skin systems and the arrangement here is that the skeleton here is no longer a hierarchy of straight rigid bones, but is rather a smooth, continuous curve. As the position of a vertex is determined by the position, radius and tangent frame of a single point on the spine, there is no need for 'weighted blending' of the influence of different bones traditionally used in skeleton/skin systems.

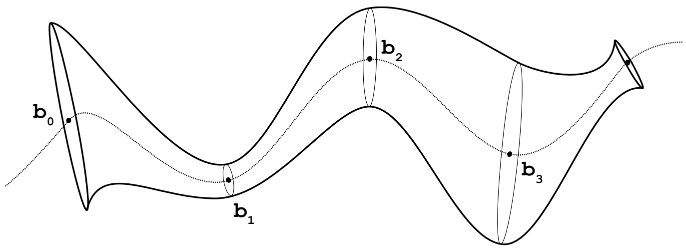


Figure 82 - The skin of the Helix.

When computing the skin, it is necessary to not only calculate the position of a vertex, but orientation also, as defined by normal, tangent and bi-tangent vectors, as these vectors are required for subsequent lighting and texturing and further animation effects. In the first implementation of the Helix, this was done analytically, using the rate of change of radius and the curvature of the spine to calculate a new normal. It was an elegant and computationally efficient solution, but it limited all subsequent animation and deformation of the skin to processes that explicitly provide the change of normal (which typically means a differentiable analytic expression). This limitation was found too restrictive, so as more powerful graphics cards became available and the number of instructions in a vertex-program was no longer a limitation, the more computationally expensive, flexible and robust *finite-difference* method was adopted. With this method, the orientation of each vertex (normal, tangent and bi-tangent), is found by computing the positions of four nearby positions on the skin, and allowing these to determine local orientation of the skin.

4.6.3 The Path of the Helix

The Helix traces a meandering path through space, at times straight or gently curved, at others spiralling, twisted and tortuous. To effect this, a simple but flexible animation framework was created that permits the layering of different animation effects. The framework is constructed around the concept of a *HelixModifier* - a self-contained algorithm that modifies some property of the spine or skin.

The principle *HelixModifier* that gives the Helix its shape does so by applying noise to the curvature of the spine. The *HelixSpineNoiseModifier* uses two 1-dimensional multi-octave Perlin noise fields[191] defined along the z-axis of *helix-space* to perturb the orientation of the bones, with one field influencing the amount and the other the direction of the bend and twist. By changing the strength, scale, lacunarity, gain and number of octaves of these noise fields, the curvature of the Helix can be made to vary from straight to sinuous to labyrinthine, on both small and large scales. Relatively-prime[192] scaling factors and phase offsets for the different noise fields ensure that repeating patterns never arise. A variable expressing the maximum permissible local curvature is also used to ensure local curvature does not exceed a certain limit, and to influence the global curvature of the spine.

The algorithm described here is an iterative stochastic process. It advances bone-by-bone along the spine, computing the orientation and position of each bone from the orientation and position of the last. The path traced by the spine is therefore a type of *random walk*; a path constructed from a succession of steps taken in random directions (Pearson, 1905). In fact, the trajectory traced by the spine is similar to the haphazard motion of a body suspended in a fluid, known commonly as Brownian

[191] Summations of multiple octaves of Perlin noise fields is referred to here as *fractional Brownian motion* - or fBm for short. Closely related, summation of the absolute value of multiple octaves of Perlin noise is referred to as *turbulence*. See Perlin (1985); Perlin and Hoffert (1989); Perlin (2002) , D. Ebert et al. (2003) and Musgrave et al. (1989).

[192] i.e with no common factors

motion or pedesis. The phenomenon is named after botanist Robert Brown, who in 1827 wondered at the haphazard motion of pollen particles floating in water under his microscope.[193] Unknown to Brown, the cause of this "spontaneous or inherent" motion had been precisely divined by Lucretius in 60 BC when, in his treatise on the atomic nature of the world, he wrote:

> Thou turn thy mind the more unto these bodies
> Which here are witnessed tumbling in the light:
> Namely, because such tumblings are a sign
> That motions also of the primal stuff
> Secret and viewless lurk beneath, behind.
> For thou wilt mark here many a speck, impelled
> By viewless blows, to change its little course,
> And beaten backwards to return again,
> Hither and thither in all directions round.
>
> Lucretius, *Of the Nature of Things*, Book II *(trans.* William Ellery Leonard)

The "viewless blows" described by Lucretius are the constant bombardment of air molecules, and in attributing the dance of a dust-mote caught in a sunbeam to unseen atomic collisions, Lucretius has perfectly explained the source of Brownian motion. Nonetheless, it would be a further 2000 years before modern physics (and no other than Einstein himself) would confirm his deductions to be true (Einstein, 1905).

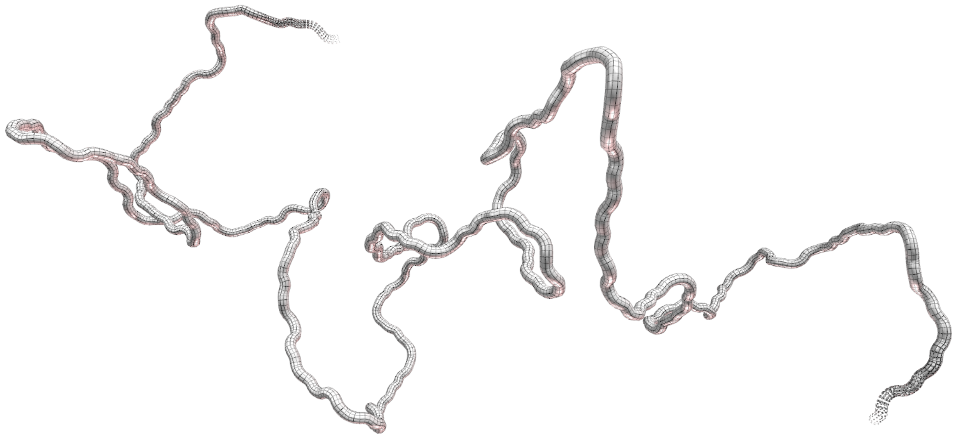


Figure 83 - Random walk of the Helix.

[193] R. Brown (1828). See Mörters and Peres (2010) for a recent account of the phenomenon of Brownian Motion.

And so it is that the Helix takes a path described so elegantly by Lucretius so long ago; a mote of dust suspended in a sunbeam. Just as on the slopes of Gargaphie,[194] where Actaeon wandered "with aimless steps"[195] through "the mazes of the pathless wood,"[196] so to the Helix traces an accidental path through space. It also describes the path of Actaeon the Stag, where the 'viewless blows' that hail down on Actaeon are not just those of the hounds at his heels and flanks, but the pulsing of his inner tumult, his torment, memories and fears, in his hopeless search for refuge.

And as much as the Helix follows the footsteps of Actaeon, man or stag, it also traces the paths of his hounds. For it has been shown that the movement of animals in the wild closely resembles pedesis. In particular, the foraging animal, an animal on the prowl in search of new quarry, is known to follow a particular kind of random walk known as *Lévy flights.*[197] A Lévy flight is much like Brownian motion - chaotic and meandering - but periodically punctuated by long, relatively true and unwavering trajectories. This is the animal abandoning one barren hunting ground for another, and it is the path taken by Actaeon's hounds as they search haplessly for their master, having long ago unwittingly devoured him. This meandering, restless flight would be the inheritance of the VøSPAZÀR, the descendants of the dogs of Actaeon, a vagabond people fated to wander like stray animals in a perpetual diaspora.

The *HelixSpineNoiseModifier,* by decoupling the direction of curvature from the amount of a curvature and having them controlled by two different stochastic distributions, exhibits Lévy flight like behaviour. There are periods of high curvature where the Helix turns knot-like upon itself are interrupted by periods of low curvature,

---

194 Gargaphie, the valley sacred to Diana where Actaeon was torn to pieces by his hounds.

195 Actaeon, "free of his share of the labour, strays with aimless steps through the strange wood, and enters the sacred grove. So the fates would have it." (Ovid & Kline, 2000).

196 "That trod the Mazes of the pathlesse Wood" (Ovid & Sandys, 1632).

197 The name "Lévy flight" was suggested by Benoît Mandelbrot in deference to mathematician Paul Lévy (1983). Viswanathan et al. (1999) argue that Lévy flights are an optimal search strategy for locating randomly distributed objects and go on to show that the movements of bees, deer, amoeba and albatross all trace such Levy-flight-like movements. To this curious menagerie can be added spider monkeys (Ramos-Fernandez et al., 2004), great white sharks (Sims et al., 2012) and apparently, humans (Rhee et al., 2011; Scafetta, 2011). The proposition that amoeba seek their prey in much the same manner as the great white shark is a wonderful image.

where it uncoils and languidly stretches out, probing tentatively the space around it, before striking out for a new hunting ground.[198]

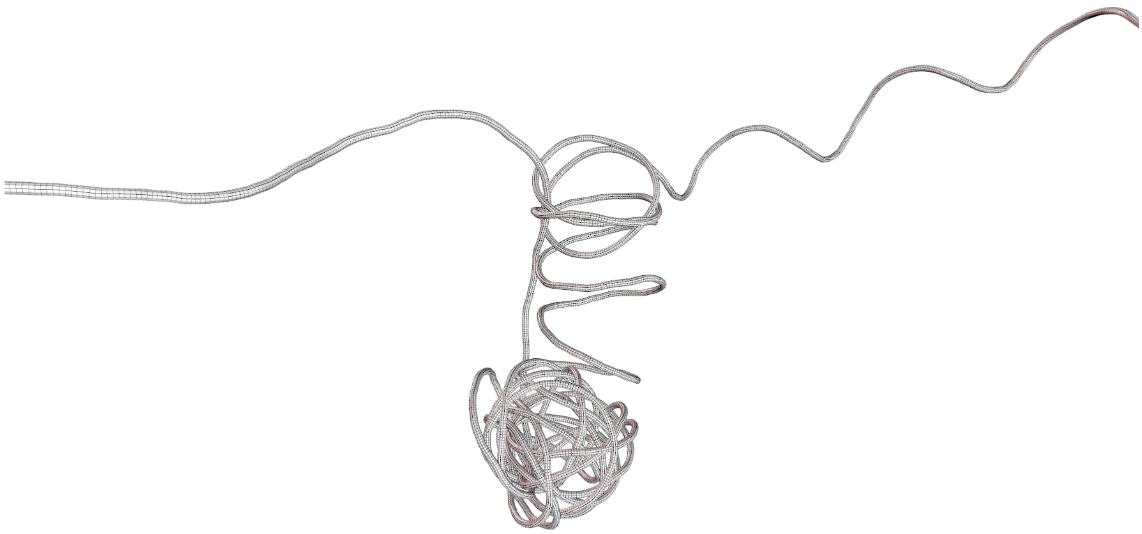


Figure 84 - The Helix following a Lévy flight-like trajectory.

When the Helix does pick up the scent, a different, more direct behaviour comes into play. The *HelixTargetModifier* provides an attractive force between the bones of the Helix and some specified point, object or region in space. Similar to the gravity field of a planet, with parameters controlling the radius, event horizon and decay of the gravitational attraction, this modifier can gently draw the Helix towards a destination, propel it like an arrow through a target, or guide it into complex and chaotic orbits. Negative forces can be used to repel the Helix rather than attract it. Despite its simplicity, the target modifier offers many possibilities, and can lead the Helix on highly complex, convolute paths. Multiple targets can exert influence simultaneously, be themselves put in motion, or even attached to the Helix itself.

[198] The path of a foraging or meandering animal is not the only thing alluded to in motions of the Helix. By modelling the Helix as a chain of elements, with angles between them following a stochastic distribution, we have emulated the shape of that other helix - DNA. For the twisted path of the DNA helix "is best described as a stochastic or random walk process, having forward, retrograde, and sidewise individual steps, but with an overall sense of bending." (Dickerson et al., 1983).

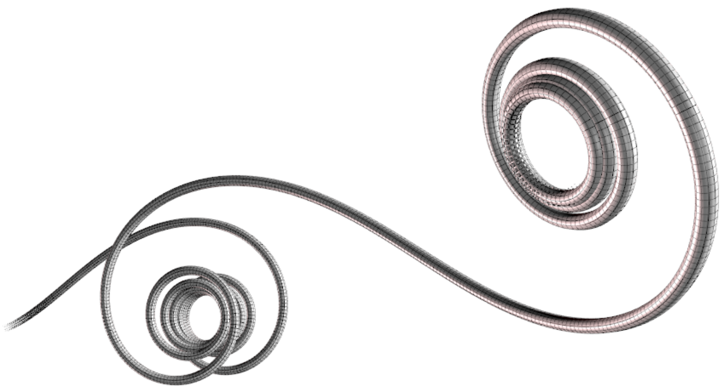


Figure 85 - Helix oscillates between two TargetModifiers, which behave like orbital traps.

4.6.4 The Shape of the Helix

A second class of Helix modifiers operate not on the curvature of the spine, but on the radius of the bones. The *HelixRadiusNoiseModifier* uses a 1-dimensional noise field to influence the radius of each bone, and by modifying the strength, scale, lacunarity, gain, number of octaves and power distribution of the noise field, a variety of forms emerge. Similarly, the *HelixRadiusWaveModifier* applies a sinusoidal pattern to the bone radii to produce regular pulses. Here the Helix recalls the harmonic resonance of a vibrating chord. The *HelixRadiusFlareModifier* flares the ends of the helix, producing two gaping orifices, not unlike the horn of a trumpet.

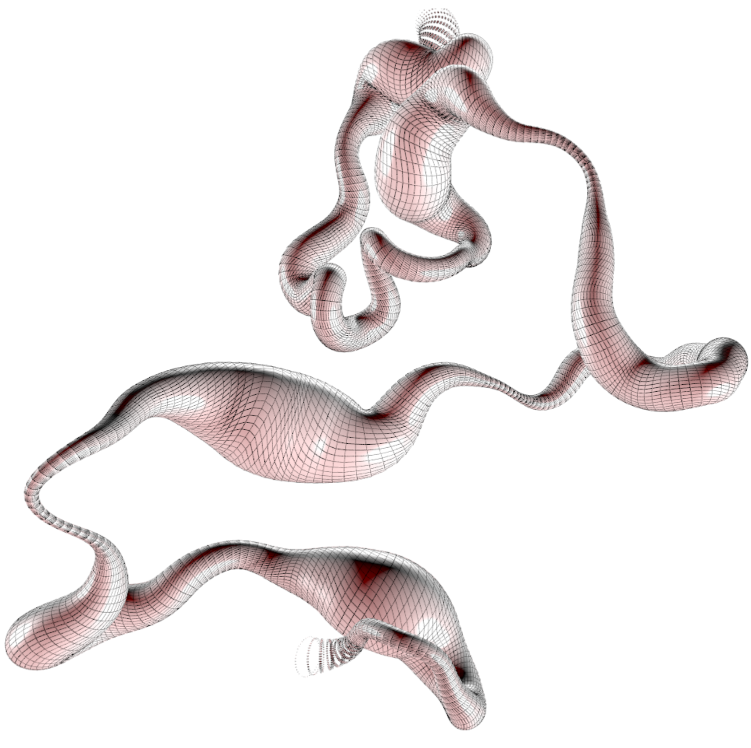


Figure 86 - Helix Radius modifiers in action.

The *HelixSkinNoiseModifier* differs from the previous modifiers introduced so far, in that it does not operate on the Helix's spine or bones, but on the skin directly. Here, a 4-dimensional fBm noise field is used to perturb the radius of each vertex. Although a 2D field is sufficient to define a scalar offset at each point on the Helix's surface (θ, z), it would result in a visible seam in θ at $0 > 2\pi$ boundary, for the values of the noise field at $\theta = 0$ and $\theta = 2\pi$ will be unequal. To avoid this, a 3D field, indexed with $(\cos(\theta), \sin(\theta), z)$, is used instead. This field is then put in motion by adding an extra dimension for time.

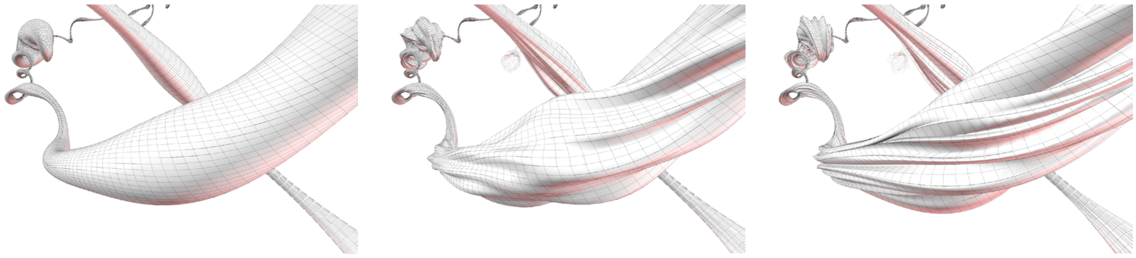


Figure 87 - Helix skin-noise modifier.

This skin noise not only introduces an asymmetry into the Helix's shape, but provides it with a carnal nature, as if an anatomy of muscles, tendons, bones and organs were labouring beneath the skin.

The combination of these skin and radius modifiers allows for the breadth of the Helix to vary dramatically. At its narrowest, the Helix might be a mere one metre wide, while at other times, when the Helix is greatly engorged, its belly might distend to form caverns hundreds of metres in diameter.

Path and Vessel

As the viewer travels the length of the Helix, bones and skin sections at the tail decay and fall away as new ones grow at the head. This is achieved graphically by having the skin at the tail, crack, fray and crumble into loose elements, which are then dispersed into the void. At the head, the opposite occurs, where a cascade of fragments fly in from afar, growing to the exact size and shape needed to assemble themselves into a new patch of skin. Like this, the Helix undergoes a continual process of growth and decay.

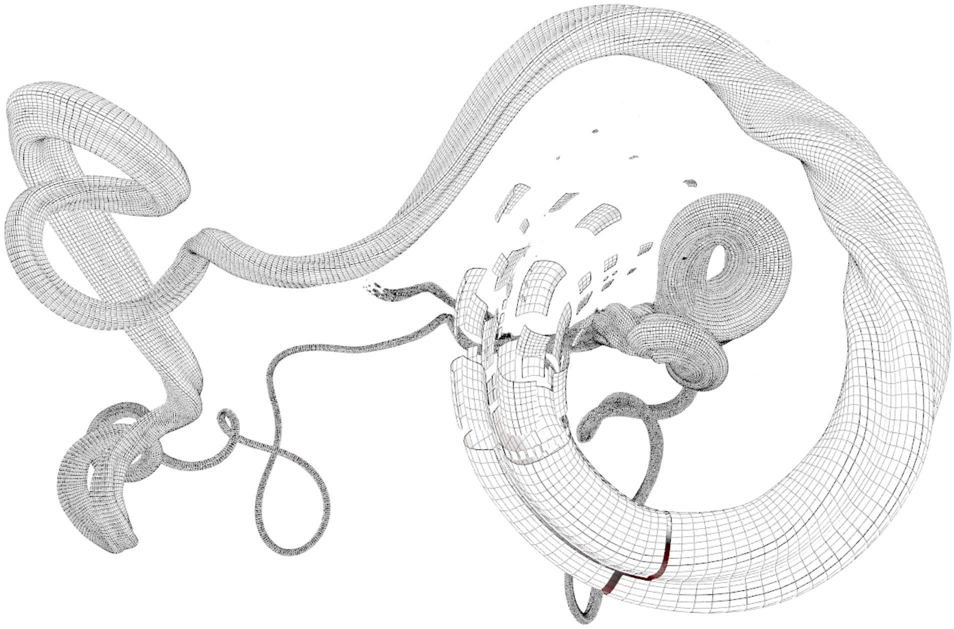


Figure 88 - Accretion of new cells at head of Helix.

The Helix is, therefore, simultaneously stationary and in motion. Even as the Helix accompanies the viewer on their journey, between the accretion of the head and dispersion of the tail, the Helix is at a stand-still. This opposition permits the Helix to play two very different roles in the unfolding narrative. As a stationary object, the Helix is a path, or conduit, along which the viewer journeys. At the same time, the Helix is a vessel conveying the viewer towards an unknown destination, or, when seen from outside, a fellow traveller.

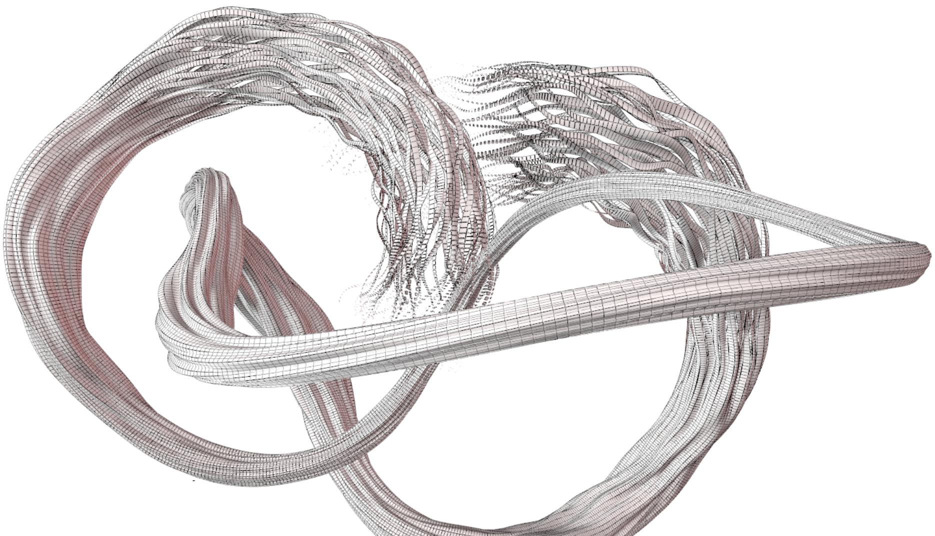


Figure 89 - Tenticular decay and growth at the head and tail of the Helix.

Local and Global Helix-space

This 'stationary motion' of the Helix is best described mathematically by distinguishing between local helix-space and global helix-space. Local helix-space is defined to be always 0 to 1 along the visible section of the spine. As the Helix advances and bones are recycles from the tail to the head of the spine, their local z-coordinates are perpetually readjusted to maintain this unit interval. In global Helix-space, however, the z-coordinate of all bones remains fixed, and each new bone grown at the head of the Helix is given a higher z-coordinate than the last. The range of z-coordinates in global helix-space advances unrelentingly. The difference between the two spaces is evident if one imagines an object with a fixed z-coordinate. A stationary object defined in local Helix-space travels along with the viewer, while a stationary object in global Helix-space is something that the viewer travels by.

The result of all this is that an effect grounded in local helix-space moves *with* the Helix as it advances on its journey. In contrast, the Helix and viewer journey *through* or *by* effects defined in global helix-space. In the case of the *HelixRadiusFlareModifier*, which is grounded in local-space, the result is two gaping orifices fixed at the beginning and end of the visible portion of the Helix. The shape of the expansion can be convex or concave, subtle or significant. At its most extreme, the ends of the Helix can be engorged so as to fill all visible space, creating the effect of two planetoids, linked by something resembling an umbilical cord.

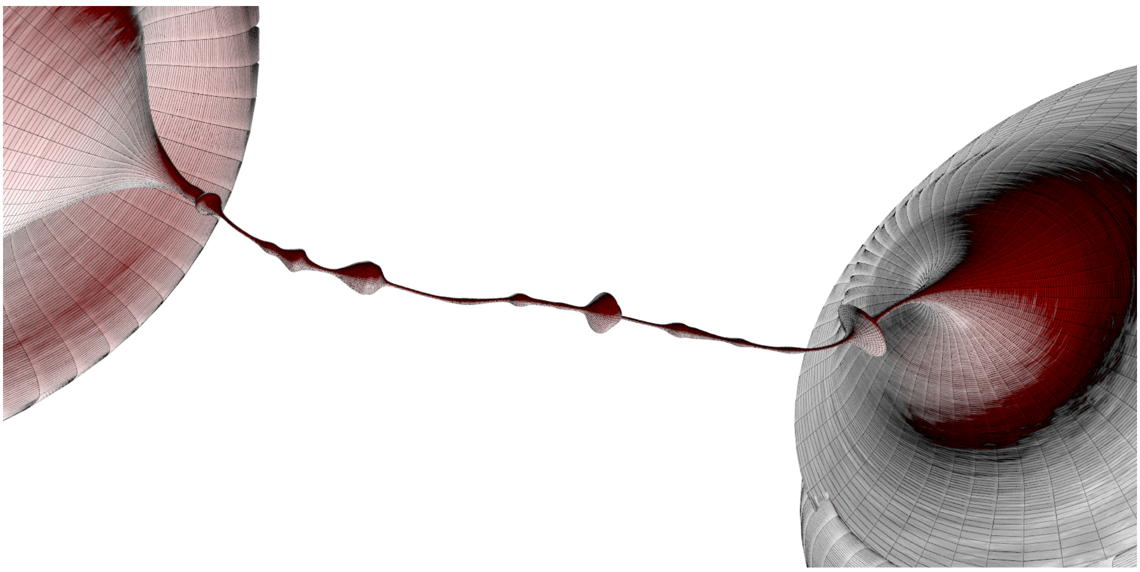


Figure 90 - HelixRadiusFlareModifier, operating in local Helix-space.

Figure 91 - Helix distended to form vast stationary surface. Still from video captured in AVIE.

As the Helix advances along the spine, and new skin accrues at the head, the *HelixGapModifier* may infrequently introduce gaps between skin sections. With parameters controlling the likelihood, length and uniformity of gaps, a variety of rhythms between gap and skin can be achieved.

Autophagy and Self-Repulsion

The *HelixSelfAttractorModifier* provides a form of attraction or repulsion between bones within the Helix. When the force between bones is repulsive, the trajectory followed by the Helix can be described as *self-avoiding*. When inactive, however, the Helix is capable of turning in on itself, and self-penetrating. When portions of the Helix are greatly distended, and others narrow, nested structures can form, in which the Helix inhabits itself.

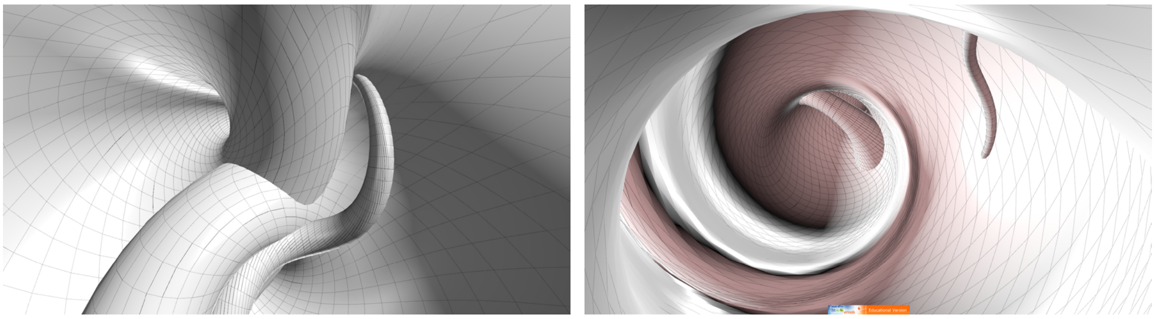


Figure 92 - Self-penetration and nested structures.

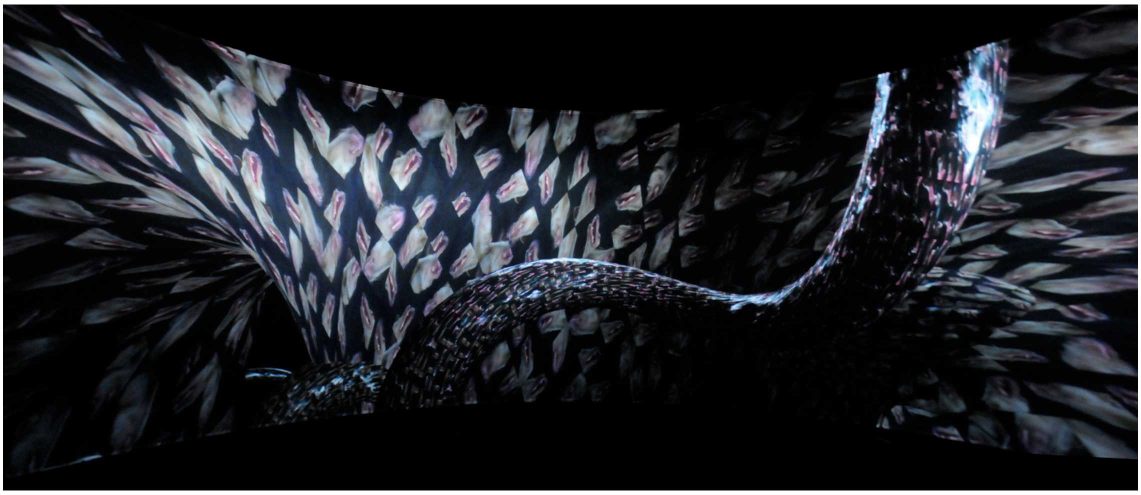


Figure 93 - Self-penetration of the Helix. Photo captured within AVIE.

When the Helix is possessed by a self-attraction, and its curvature and twisting is just so such that the Helix curls back upon itself, the Helix can appear to devour itself. As the particles of the tail are shed like dead cells into the mouth of the Helix for renewal, this act resembles *autophagia*. It is at these moments that the Helix resembles the *Ouroboros*, the self-devouring snake heralding, for alchemists, cyclicality and eternal rebirth.

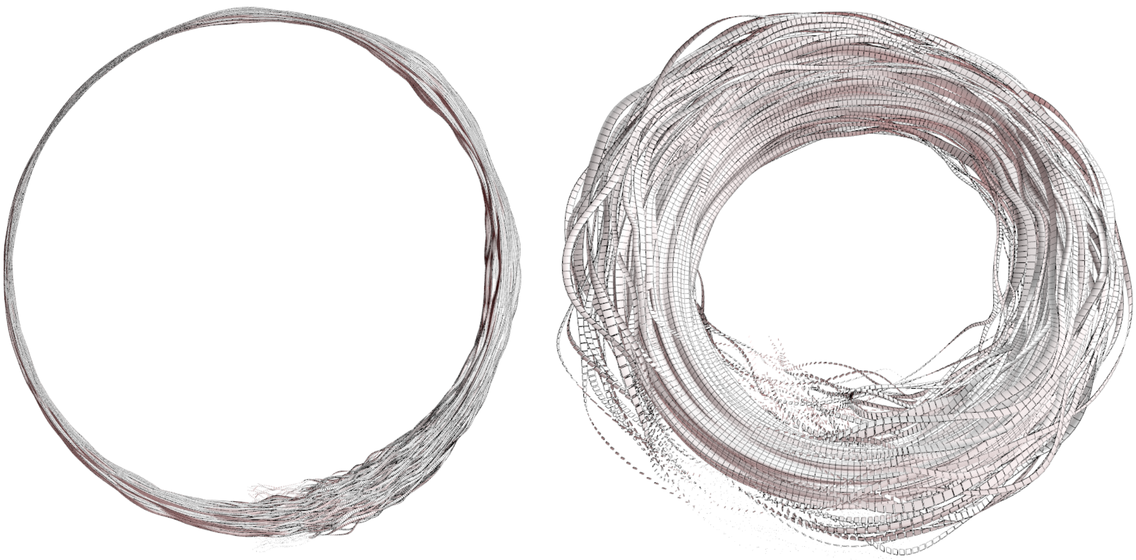


Figure 94 - The Helix in varying states of autophagia.

4.6.5 Animating the Helix

By altering the parameters of the above modifiers, and the parameters of their associated noise-fields, the Helix is put in motion. Parameters can be altered gradually, inducing change that is imperceptible (or perceptible only in retrospect) or abruptly, with force or violence. Or cyclically, to produce rhythmic or oscillatory motions. They can be altered in unison, or individually. Noise-fields can be given velocity, which for those modifiers that operate in Helix-space, causes the effect - a bulge, ripple, coil or kink - to travel along the length of the Helix or twist around the spine.

By setting it in motion, the Helix is brought to life. Motion - in the absence of obvious external causes - is one of the quintessential tell-tale signatures of life. Indeed, before the age of autonomous machines, self-motion and life were held to be synonymous. To be animated was to be alive, or as the Spanish Jesuit Toletus wrote in his commentaries on Aristotle:

> To live is nothing other than a certain self-motion, and movement of one's own is a kind of life: and all things which by some motion move on account of themselves, and by an essential principle… live. Life consists in this, that something essentially has an active principle… of some motion within itself.
>
> Toletus, *Octo Libros Aristotelis de Physica*, 1572[199]

Such is our natural inclination to perceive spontaneous motion as evidence of life, that when Robert Brown perceived those twitching particles of pollen through his microscope, his first assumption was that they were alive. Even after observing the mysterious motion in particles obtained from all manner of inorganic materials, such as metals, rocks, coal, glass, and even "a fragment of the Sphinx" (R. Brown, 1828, p. 472), and despite all his better knowledge, he could not help but see the particles as animated *animalcular*. Some twelve years prior, James Drummond had observed similar microscopic motions in the eyes of dead fish, and could not resist drawing the same conclusions as Brown:

> Perhaps many other objections may be opposed to the supposition of animalcular life in these bodies; and yet the strong *expression* of animation, if I may so term it, and air of seeming design, with which the varying motions,

[199] book 2, chapter 2, text 15, q. 2 "*Vivere non aliud est quam quoddam se movere, et vita quaedam motus sui: et omnia quae se aliquo motu per se, et a principio essentiali... movent, vivunt.... Vita consistat in hoc quod aliquid habeat essentiale principium ... activum alicuius motus in se.*" Translated in Byers (2006)

> sometimes slow, and sometimes rapid, are performed, and the difficulty of otherwise accounting for their motion, whether real or apparent, lead, upon the whole, I think, to this supposition, not as one which we can admit with confidence, but as the *least improbable* conjecture, which, in the present limited state of our knowledge, we can venture to form. (Drummond, 1815, p. 384)

What drew both Brown and Drummond towards this unlikely conclusion, and against their modern understanding of the world, was the nature of the motion of itself. It was "nothing like attraction, or repulsion," but "proper to itself", with a "writhing and twisting" that "might be said to be somewhat vermicular," such that "to account for the motions, the least improbable conjecture is to suppose the spicula animated."

It would seem that Brownian motion is capable of triggering *biological motion* detection, a perceptual faculty specifically dedicated to the detection and recognition of animals in motion. Since Johansson (1973) revealed the curious human capacity to perceive and recognise all manner of human motion in visual stimuli composed of nothing more than a dozen moving points, a great deal of evidence has been amassed indicating that the perception of biologically induced motion is an innate and intrinsic capacity of the visual system.[200] Our sensibility to biological motion is incredibly nuanced, for we are able to distinguish between animals, gender, actions, facial expressions and even emotion by observing nothing more than a dozen moving dots in motion. This sensibility is explained by supposing that the visual system is furnished with dedicated organic-motion detectors, equipped with their own special neural pathways (Ptito et al., 2003, Kaiser, 2012 #151). Troje and Westhoff (2006) describes these "life-detectors" as visual filters tuned to the "invariants of animal locomotion," and as these filters exists in both newly hatched chicks and humans (Simion et al., 2008), they are most likely an "evolutionary old mechanism that the human visual system shares with other animals."[201]

---

[200] See Blake and Shiffrar (2007) for an overview. See Mather and West (1993) and Bellefeuille and Faubert (1998) for discussion specifically related to the perception of animals in motion.

[201] The author, who suffers from arachnophobia, often finds that spider-like motion is sufficient to trigger a physical repulsion.

The organic character of motion is, however, just one aspect of the perception of animacy. Shape, structure and the relationship between the agent and its surroundings also play a role (Scholl & Tremoulet, 2000). One factor is the appearance of *self-motion.* Aristotle, among others, considered self-motion to be *the* singular defining aspect of life, echoed here in the Aristotelian commentary of the Conimbricenses:

> That thing whose movement is from the outside, is inanimate, but that to which it is intrinsic to itself to be moved by itself, is alive. [...] To move oneself immediately and as the principal cause of one's own motion is the proper office of life. (Commentarii Conimbricenses, *Octo Libros Physicorum Aristotelis*, 1592, 1602).[202]

The perception of self-motion is a form of causal perception - it is the perception of an event without a visible cause. It is, therefore, a direct and uninterpreted perceptual faculty. Both the perception of self-motion and organic motion can be expected to play a part in the perception of animacy, although their relationship is likely complex. For one, it is possible to perceive, to a certain degree, animacy in the absence of self-motion, as is evidenced by puppetry. In contrast, it is common-place to perceive self-motion with perceiving life; a motor car being one example. Further, perception of animacy has been show to influence perception of causality. For example, the permissible distance and delay between cause and effect depends on whether any of the entities are perceived as animated or not (Falmier & Young, 2008). Despite this interplay, there is evidence that it is the quality and trajectory of motion that plays the dominant role in the perception of life, with features such as self-motion, shape and structure offering supporting roles (Bassili, 1976; Berry et al., 1992; Scholl & Tremoulet, 2000).

Perception of animacy is, therefore, not unlike the perception of causality, in that they both involve the direct perception of phenomena that we might ordinarily imagine to be the product of inference or reasoning. That is, one might expect that we infer that an entity is alive from its motion. This, however, is not the case. We physically and

---

[202] Book 8, chapter 3, q.1 a.2. "*cui forinsecus est moveri, id inanimatum est, cui vero intrinsecus sibi ex se, animatum*", "*movere se simpliciter et ut principalem sui motus causam est proprium munus vitae.*" Translated in Byers (2006).

directly perceive animated life in just the same manner we directly perceive colour or shape. Certainly, "high level" cognitive reasoning and knowledge is known to influence this perception, just there are equally high level forces in the perception of causality, but the dominant force at play is this low-level, physical, inescapable, uninterpreted and unconscious direct perception. Drummond and Brown, despite knowing better, were no more able to resist seeing the dancing particles as living *animalcular* than one is able to resist seeing the colour red.

To summarise: under certain conditions, fractal Brownian motion may appear life-like, and this perception is a low-level, immediate and irresistible response, rather than a high-level, cognitive act of reason, inference or interpretation. This appears to extend well beyond pure fractal Brownian motion and random walks to all manner of motions that share the same stochastic basis - be it the movement of joints or appendages, or the undulation of a surface. This motion appears to fit some learned or innate invariant template or "visual filter" normally reserved for the perception of organic motion.

Admittedly, this leaves an explanatory gap, for while it describes what is happening, it does nothing to explain why this particular mathematical model describes so many natural, organic phenomena so well. Enriquez (2004, p. 1) suggests that the apparent naturalness of fractal Brownian motion is due to three characteristic features: "it is a continuous Gaussian process, it is self-similar, and it has stationary increments." It would seem that Enriquez is implying that these are features common to organic processes, which may be true, but it begs more questions than it answers. Such questions are beyond the limits of this thesis, and certainly beyond the limits of my knowledge. Instead, I will have to remain satisfied with the simple observation that Brownian motion and similar stochastic, self-similar motions are effective tools for triggering the perception of life, and the life-like aspect of these motions is as much a question of perception as it is mathematics.

From animat to animal

By animating the Helix in a manner that is perceived as natural or organic, the Helix is transformed from a structure into a living form. The perception of internal organic function and motion is irresistible. Oscillations and cyclic motion of parameters produce rhythmic behaviours evocative of throbbing, twitching, pulsing or breathing. Travelling bulges and lumps appear like peristalsis. Movement appears anatomical, and one can almost see muscles and tendons at work.

Once an entity is perceived as being animate, attribution of volition and intentionality, mood and mind follow soon after. As the seminal experiments of Heider and Simmel (1944) demonstrated with their simple animated blocks, (and countless studies after), the perception of motion as intelligent, intentional or emotional can be elicited with the barest of stimuli. Despite its alien form, the invitation to anthropomorphise (or at least zoomorphise) the Helix, and perceive in its movements the expression of sentience and will, is difficult to resist.

By animating the parameters of motion over time, the Helix's behaviour ranges from languid to frenzied. In doing this, the Helix is given a personality. It appears to be in possession of different moods, or rather, to be possessed by different moods. The Helix's transformation from structure to animal to the principle protagonist/antagonist is now complete.

It is important to note that all of these perceived organic behaviours and outward signs of a mind or personality, are *entirely emergent*. They are not designed, and no explicit model of mood or anatomical feature exists in the code. The parameters are simply randomised, and these phenomena exist entirely in the eyes of the observer.

4.6.6 Inertia

Animating the parameters does however present some challenges. For example, when altering the scale of a noise-field, changes around the origin of the field are small, but as distance from the origin grows, the more violent the changes become. At some distance from the origin, no matter how slowly the scale is adjusted, the resulting changes in the value of noise-field are too rapid to be perceived as coherent.

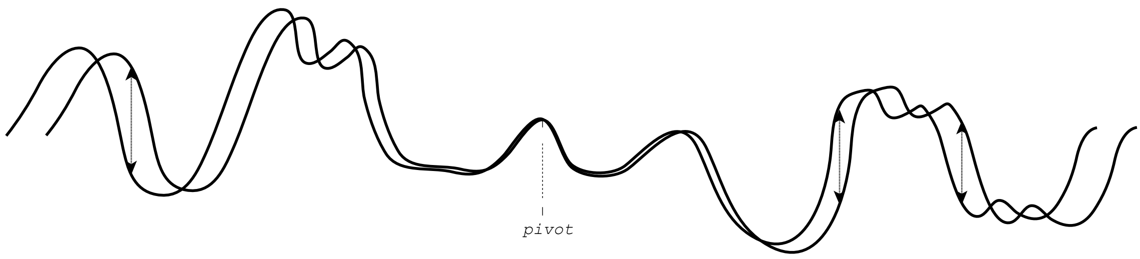


Figure 95 - Scaling of noise fields. When the horizontal scale of a field is changed, the amount of local change is a function of distance from the pivot of the scale.

One remedy for this is to pivot the scale around the viewer's position. When changing scale, we simultaneously adjust the phase of the noise-field, so that the rate of change of the noise-field at the viewer's position is zero, but increases with distance from the viewer. That is, for a given pivot position x, we wish the value of the noise field to be constant, regardless of the scale s and phase p: i.e. $f(sx + p) = f((s + \Delta s)x + p)$. This is easily achieved by adjusting the phase by the amount $z = - x\,\Delta s$.

If

$$f(xs + p) = f(x(s + \Delta s) + p + z)$$

then

$$z = -x\Delta s$$

where f is the noise-field,
s is scale, p is phase, x is position,
Δs is change in scale,
and z is the offset applied to effect a pivot.

Eq. 29 Pivoted Noise field

Like this, changes in the noise-field caused by a change of scale are less dramatic near the viewer, and more chaotic at the ends of the Helix. This technique can also be used on multi-octave noise functions (fBm, turbulence, ridged-multifractal) that use a different scale for each octave; we need only calculate and store a unique phase-offset for each octave.

A similar but perhaps more profound problem exists when animating the parameters of the Helix modifiers that operate iteratively over the bones, such as the spine-modifiers *HelixSpineNoiseModifier* and *HelixTargetModifier*. Because the position of each bone is calculated from the position and orientation of the bone that precedes it, small changes at the beginning of the spine can accumulate rapidly into large changes in position at the other end. In fact, the system is at times chaotic, in which case no matter how slowly and fine-stepped the parameters are altered, unacceptably large jumps in position of the end of the spine may occur.

To alleviate this problem, the Helix is given inertia. A limit on how fast each bone can travel is imposed, and positions and orientations of the bones are calculated as a blend of the new proposed state and the old state.

These measures prove effective, but are no more than make-shift solutions to a deeper underlying problem, which is that there may not always be a continuous path through state-space between two 'interesting' or 'acceptable' states that does not pass through 'uninteresting' or 'unacceptable' states on the way. This, I posit, is a fundamental problem facing all kinds of generative and procedural systems that involve some form of exploration of parameter or 'genomic' space. If one hopes to discover truly novel, previously unimagined states, one cannot ahead of time guarantee that all of state-space is well-behaved. All art requires an element of risk, just as any voyage of discovery demands the unknown.

In summary, taken alone, each of these modifiers produce relatively simple behaviour. It is their combination that yields a surprising variety of phenomena. By using simple algorithms that have been shown to mimic biological phenomena and processes - fractal noise and random walks - and by animating all the parameters that determine the shape and path of the Helix, the Helix is brought to life.

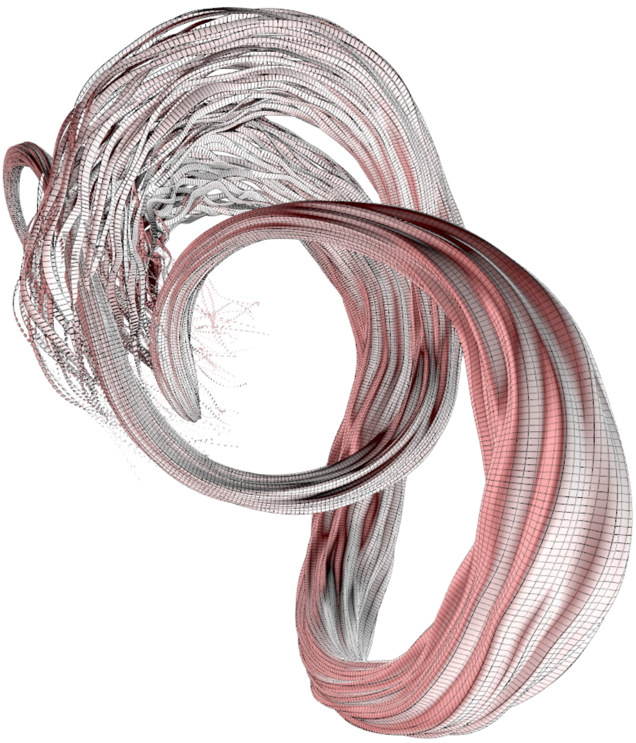


Figure 96 - The Helix as animal.

4.7 LIGHT AND SURFACE

Having discussed the form and movement of the Helix, discussion can now turn to the question of its appearance. For this, we must return to notion of ecological optics.

4.7.1 Ecological Optics and Visual Realism

It was argued in Section 2.3.3 that any act of immersive mimesis must define a form of ecological optics[203] - a thorough nomological and algorithmic account of how the virtual world manifests ultimately as physical light. It was reasoned that, if the viewer is to experience a sense of presence when perceiving the world, then this algorithmic transformation of the virtual world into light must possess certain qualities, or obey certain rules, and these rules and qualities are dictated not by the world, but by our biology. However, the principal of physical mimesis suggests that by simulating the physics of the real world, presence is assured, even in the absence of a full understanding of human perception. This was considered the *brute-force approach* to immersion. A more nuanced approach was that of metameric mimesis, in which the limits and acuity of our sensory apparatus are exploited to simplify the physical simulation, ideally yielding a vision of the world that is perceptually identical to a true physical simulation.

However, even when metameric approximations are taken into account, the computational demands of real-world optics remain beyond the capabilities of current real-time rendering systems. Fortunately, it seems that for presence to survive, the virtual optics need only remain partially or approximately faithful to its real-world counterpart, and there may exist some essential 'subset' of visual perception that is necessary or sufficient for presence. This idea was introduced previously as the *presence hypothesis.*

This search for partial or approximate simulations of real-world optics lies at the heart of computer graphics. Computer graphics is essentially the search for short-cuts; the optimal utilisation of limited computing resources to simulate a physics that is by

[203] By 'optics,' I mean any algorithmic description of the interaction between light and matter.

nature unlimited in computational density/complexity. However, the development of computer graphics has seldom been *explicitly* concerned with presence. More often, it is the pursuit of realism, or more accurately, 'photo-realism' that drives advances in the field,[204] while in other cases clarity and ease of comprehension, or aesthetics, might be the motivating force.

Explicit studies of the qualities of a virtual optics that might be necessary or sufficient for presence are few and far between, and what few there are provide contradictory results.[205] For example, Hendrix and Barfield (1996) measured the correlation between presence and visual detail and realism, finding both contribute to greater presence. Similarly, Slater et al. (1995) compared the influence of shadows on presence. When comparing dynamic shadows, static shadows and the complete absence of shadows, they found dynamic shadows yielded greater presence than static ones, and static shadows more than none whatsoever. However, when Mania and Robinson (2004) measured the effect of different shading models on presence, they found no correlation at all. Likewise Zimmons and Panter (2003) found no difference in reported presence between wire-frame, untextured, textured and shaded visuals. More recently, Slater et al. (2009) recorded a very loose correspondence between the simulation of shadows and reflections and presence: of 24 subjects, 17 reported an increase in presence, while 7 suffered a diminishment in presence.

In a more recent study, Insu et al. (2012) compare three increasingly realistic lighting models. In this work, Insu, Slater and colleagues have adopted a more nuanced definition of presence, dividing it in to two different sub-types. *Place illusion* is the illusion of "being there," while *plausibility illusion* is the illusion that the virtual world and events are real (Slater, 2009; Slater et al., 2010). Place illusion, they suggest, is entirely a function of the sensorimotor contingencies provided by the mediating system.

---

[204] The term 'photo-realism' reflects how computer graphics is essentially a form of synthetic photography; its aim is not to simulate direct vision of a virtual world, but rather to simulate the taking of a photograph of a virtual world.

[205] On the other hand, it could be argued that the computer game industry is implicitly concerned with the relationship between computer graphics and presence and, as such, represents a vast reservoir of knowledge and experimentation on the subject.

As these contingencies are effectively the degrees of freedom introduced above (Section 2.5.1), it can be argued that this concept of place illusion closely resembles the definition of presence adopted here. Plausibility illusion, they suggest, is a measure of the virtual world's responsiveness to the viewer's actions, and how well these responses mimic the real world. In their study, they find that the varying levels of realism in the simulation of light affected the plausibility illusion, but had no effect on place illusion.

In truth, very little can be taken from these studies, apart from an appreciation of the difficulty of studying the relationship between visual realism and presence. First, none of the experiments provide any reasons to assume that the results are universal and might apply beyond the very limited artificial worlds employed in the studies. It is more likely that the role any particular visual cue plays in presence is dictated by the nature of the scene itself. Second, in perhaps all but the most recent of the studies, none are actually comparing realism with non-realism. Rather, they are comparing two very unrealistic simulations, each with different kinds of faults, and again, the importance of these faults will most likely be contextual. Neither the complete absence of shadows nor the drawing of a single, sharp shadow are realistic simulations of light, for example, and which is more detrimental to presence is surely completely contextual. Third, the choice of variables in these experiments - different forms of shadows, lighting models or reflections - are entirely dictated by trends in graphics technology rather than being grounded in any theory of presence or perception, or connected meaningfully with the context of the virtual world itself.

4.7.2 Avoiding Conflict

Most importantly, none of these studies seem to make any distinction between the presence and absence of visual cues and the presence or absence of visual *conflict*s. According to the theory of presence adopted here, it is conflict between stimuli rather than the absence of stimuli that leads to a diminishment of presence. For example, a simple texture map,[206] be it captured photographically or handcrafted, captures within it the appearance of a surface illuminated by a particular environment and seen from a

[206] Here, I mean an image used to modulate the colour of diffusely reflected light.

particular point of view. When used in a virtual world, the texture map presents an array of visual cues that more often than not conflict with the rest of the visual scene. As the viewer moves about, the texture does more to reveal the flatness of a surface than it does simulate relief, and any lighting captured within the texture will almost always conflict with the passage of light throughout the rest of the scene, or with the light embedded within other textures.

Now, with this particular example, there are many methods for reducing or eliminating these discrepancies. More sophisticated texture-mapping strategies, in which the interaction of light and surface is more accurately simulated, such as per-pixel displacement of the surface height and normal, reflections and refractions, ambient occlusion and sub-surface scattering and so on, can be employed to reduce conflicting cues. The key observation offered here, however, is that a new graphical technique, while introducing or improving some visual cues, may also introduce conflicting cues, and the impact of these conflicting cues on presence is not necessarily the same as it is on visual realism. And it is useful to always keep in mind that a virtual untextured surface will always more accurately simulate a real untextured surface than a virtual textured surface can a real textured surface. Texture maps may make a scene more recognisable, but they do not necessarily increase immersion, and may even diminish it.

4.7.3 The Ecological Optics of *La Dispersion Du Fils*

The universe of *La Dispersion Du Fils* is rendered visible as follows:

Only light is visible, and nothing else. The world makes itself known by perverting the passage of light. There is no matter, only surface. Surfaces perturb a field of light by reflecting, absorbing, refracting, transmitting or occluding rays of light; interactions that, by altering the path, spectra or intensity of light, reveal the existence of surfaces and the structure of the world to an observer.

Within *La Dispersion Du Fils*, the interaction of light and surface is intended to crudely mimic the optics of the real world. Following Phong (1975), the reflected component of light is divided into diffuse and specular components. Diffuse reflectance represents light that is refracted into the surface, scattered, partially absorbed and re-emitted. It is this absorption that alters the colour of reflected light, and is therefore responsible for giving a surface its colour. Specular reflection is the mirror-like reflectance of light. For many materials, light undergoing specular reflection is largely untainted, although this state-of-affairs varies with both material and angles of reflection.

In the early implementations of the Helix, simple Lambert reflection was used for the diffuse component, while Blinn-Phong was used for specular (Blinn, 1977). In the most recent versions, the diffuse component was replaced by the more physically accurate (and more computationally expensive) Oren-Nayar model, in which the surface is modelled as a micro-facetted Lambertian surface, where the facets may mask, shadow or inter-reflect light (Oren & Nayar, 1994). While basic Lambertian reflectance is parameter-less, Oren-Nayar has a single measure of roughness, which can be altered as desired.

Bump mapping

In 1978, Blinn introduced a method for simulating the interplay of light and fine surface structure by altering the normal of a surface while calculating the reflectance of light (Blinn, 1978). By using a bitmap to store deviations of surface normal or surface height, the depiction of surface detail far finer than the resolution of the underlying mesh geometry can be achieved. The technique represents a decoupling of shape information into two structures: macro-scale structure is captured in polygonal meshes, while micro-scale detail is modelled as local deviations in surface orientation or height. Importantly, the resolution of the two representations are completely independent, and the technique is very well accommodated by modern GPU.

This technique is used for the shading of many surfaces within *La Dispersion Du Fils*, including the skin of the Helix. The 'bump-map' – the bitmap describing the deviations in surface height and normal at each point on the surface - was derived from the antler of a deer. A photograph of the antler was cropped and manipulated to tile seamlessly, and then used to algorithmically procure normal, height, occlusion and specular maps.[207]

Figure 97 - Antler of a deer (left) used as source for the bump map (right).

The resolution of this bump-map is 1024x1024 pixels. The bump-map is wrapped around the circumference of the Helix, which may vary from 6 to 600 metres as the Helix swells and constricts over time. As the viewer's proximity to the skin of the Helix may range anywhere from as distant as 1km, to as close as 10cm, the bump-map has insufficient resolution to meet this enormous range in scale and viewing distances. Stretched over the surface of the Helix when it is fully distended, each bump-map pixel might occupy as much as half a metre and when viewed from nearby, a single bump-pixel might occupy the entire field of view. Supposing a desired resolution of 1cm per pixel, a bump-map of 60000 x 60000 pixels would be required. This is not only well beyond the limits of the GPU, but there is insufficient information in the original photo of the antler to fill these pixels.

The solution adopted here is to apply two bump-maps at vastly different scales. One bump-map is stretched to wrap once around the belly of the Helix and provides high-level structure, while the other is applied at 1/40th of the scale, and provides micro-level structure.

This technique is very effective at adding texture to a surface. As the deviated normals are used to light the surface, shading and reflections are coherent with the

[207] The commercial tool *Crazybump* was used to compute this data from the photograph.

lighting in the rest of the scene, even with moving geometry, viewer or lights. This technique, however, has a significant flaw. Bump-mapping, as devised by Blinn, only affects the shading of a surface, and has no effect on its shape. As a consequence, a battery of conflicting visual cues are introduced. The silhouette of a bump-mapped object remains completely unchanged, as do any other texture maps applied to a bump-mapped surface. When viewed from an oblique angle, lack of self-occlusion becomes apparent, as does lack of motion parallax when the viewpoint or surface itself is put in motion. Shadows, both cast by the object and falling upon the object, also suffer from lack of deviation. Bump-mapping, it seems, is a fine example of the notion discussed above, in which the introduction of new visual cues can bring with it an assortment of new visual conflicts.

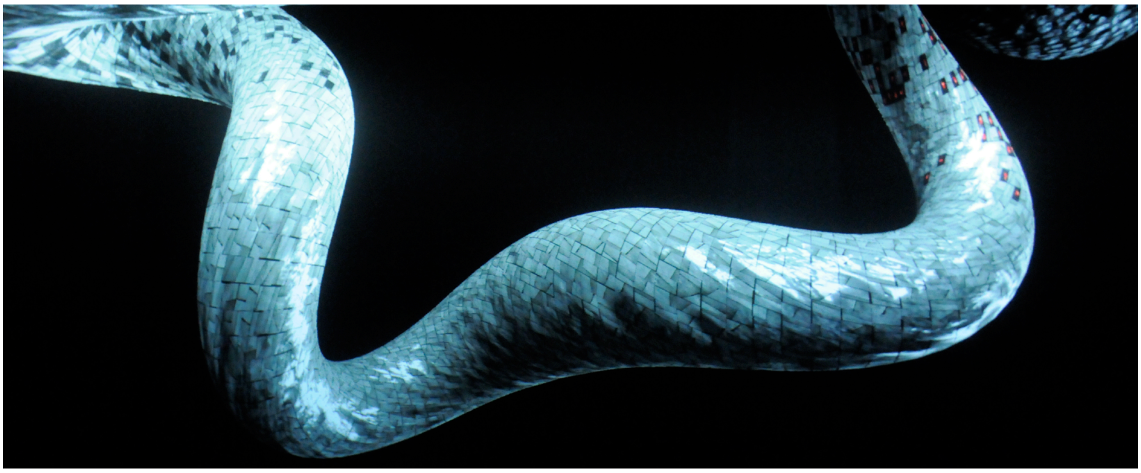


Figure 98 - The Helix with bump-mapping. Photo taken in AVIE.

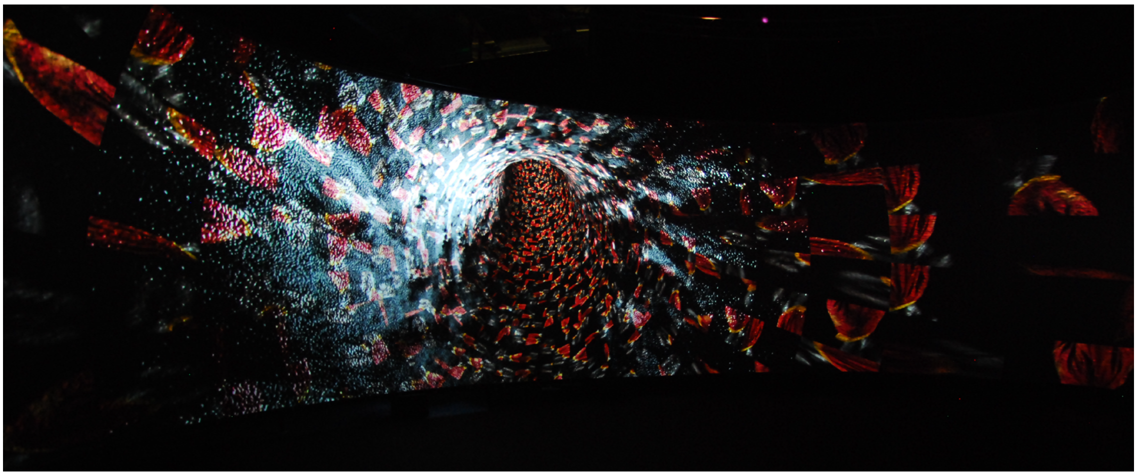


Figure 99 - The Helix interior with exaggerated bump-mapping. Photo taken in AVIE.

For the most part, when the bump-mapping is applied conservatively, most of these flaws can be ignored. For the Helix, the most significant flaw is the lack of distortion in the movie-images painted on its surface. For while the bump-maps perturb the colour and intensity of reflected light, they have no effect on the forms within the movie-images. The result is that the skin of the Helix appears to be composed of two membranes, one transparent and wrinkled and the other completely smooth. The effect, when viewed in stereo in the AVIE, is uncanny.

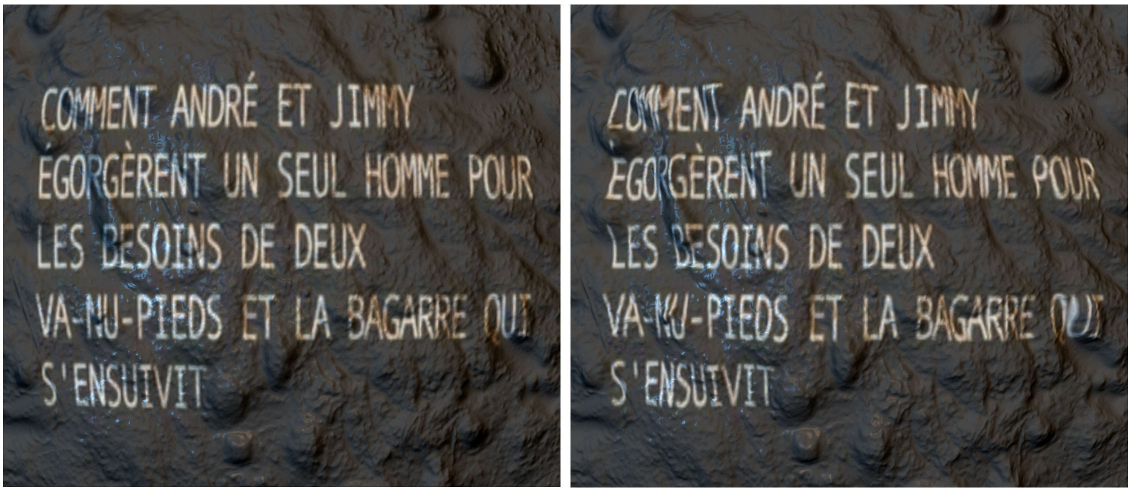


Figure 100 - Bump-mapping and Parallax mapping. Left: The normal-map alters the surface normals only, leaving the shape of the text undistorted.Right: The normal and height-map are used to offset the textures coordinates of the text, adding suitable distortion.

There are a number of improvements to Blinn's original method that alleviate some or all of these cues, at greater computational cost, by tracing a ray through the height-map to calculate the true, perturbed coordinate of the texture-map. The reader is referred to Szirmay-Kalos and Umenhoffer (2008) for a thorough review of a variety of such techniques. In the most recent versions of *La Dispersion Du Fils*, a technique known as *iterative parallax mapping* (Premecz, 2006) was implemented in an attempt to remedy this visual conflict. The results were very effective, save for one small, yet significant 'logistical' problem. By correctly simulating the height of the surface when calculating the texture coordinates, it is now possible that an image mapped to one polygon is now visible on a neighbouring polygon. Given the manner in which movies are arranged on the skin of the Helix, and the limited resources available to a single shader (3.0) program, this presents a number of technical obstacles. For example, where polygons previously required only one material, they now need nine, exceeding the number of texture-samplers that can be accessed in a single shader-pass.

Sources of Light

Within *La Dispersion Du Fils* there are a number of different sources of light. The primary source is a simple arrangement of omnidirectional point and 'directional' light sources, both popular techniques in real-time computer graphics for no reason other than their computational efficiency. For the most part two directional lights, and three point sources are at work. The directional lights work in tandem to emulate the theatrical conceit of a key-light and back-light. The three point sources provide additional complexity to specular highlights. All properties of the lights – colour, intensity, position or orientation – can be animated and adjusted to the occasion.

In addition, surfaces themselves may emit light, which is modelled very simply by adding a constant factor to diffuse reflection.

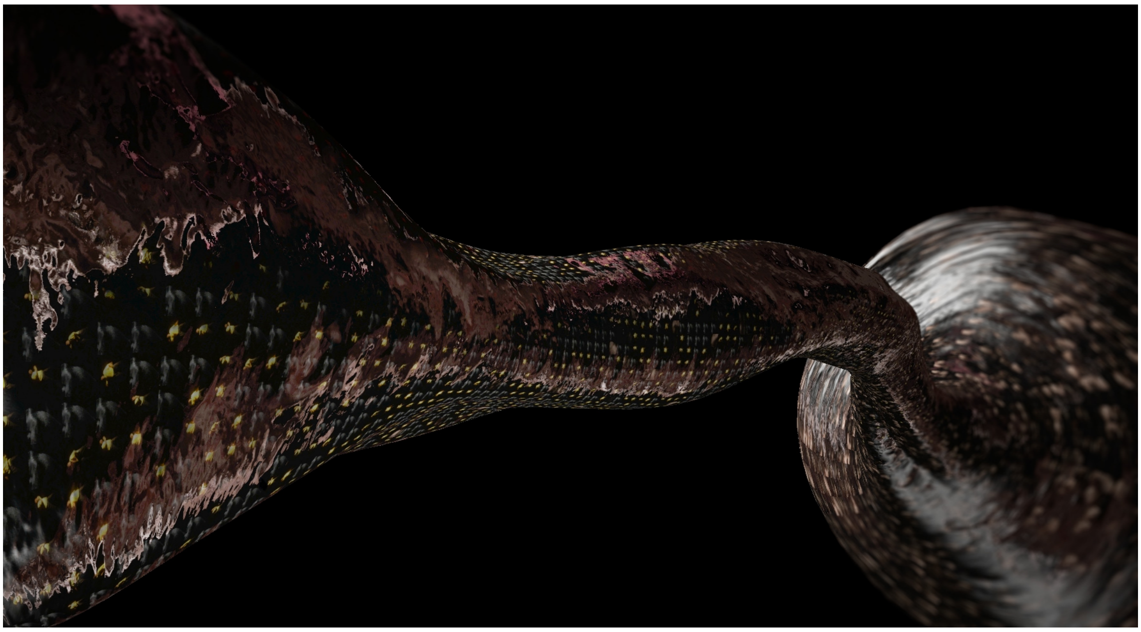


Figure 101 - Movie used as environment map.

A further source of light is the use of environment maps (Blinn & Newell, 1976). Here, images are used to define the incoming radiance and colour. The technique works very well for highly reflective surfaces, where each point on an illuminated surface need only sample a small region of the environment map. For more diffuse reflections, however, each point on an illuminated surface should ideally be lit by a large region of the environment map, necessitating a computationally expensive convolution. This is approximated by using the MIP-map chain of the environment map to approximate these convolutions, in a technique proposed by McGuire et al. (2013). This technique is highly suited to real-time graphics, as MIP-map chains can be computed very rapidly every frame, allowing rendered images or videos to serve as sources of light.

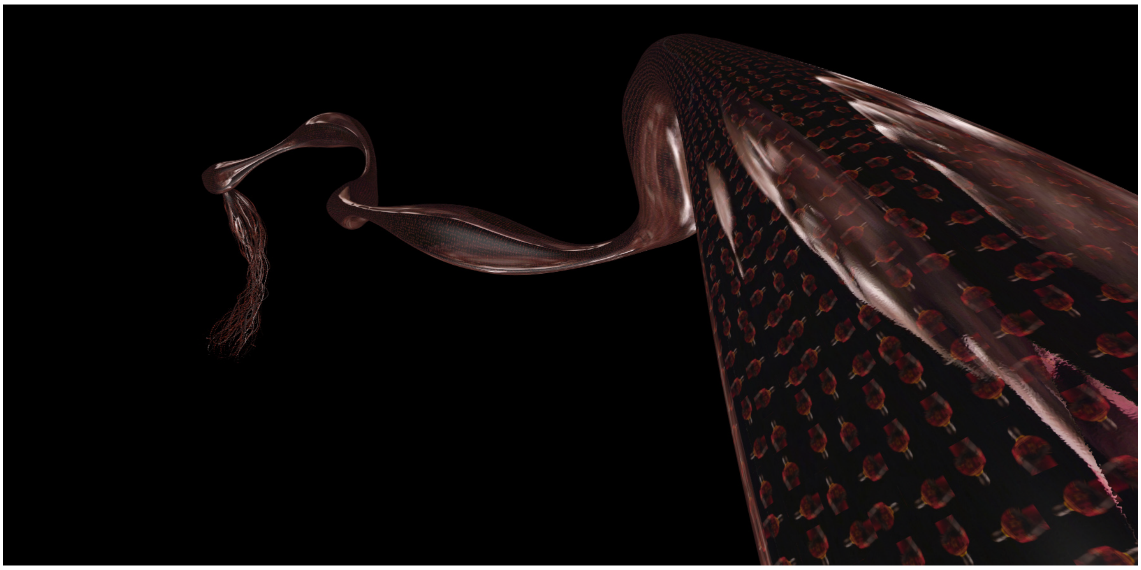


Figure 102 - Reflection mapping of movies 1.

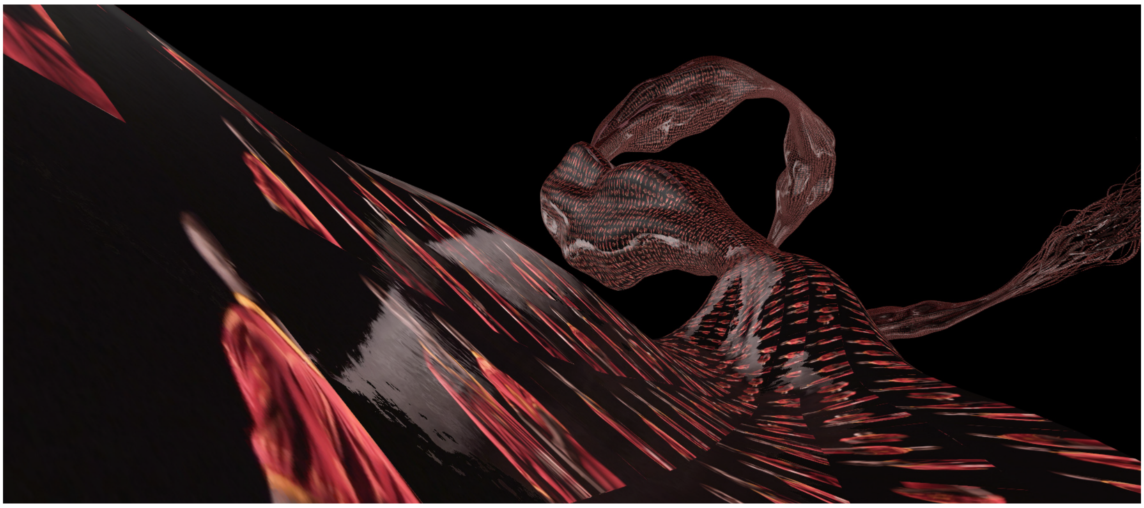


Figure 103 - Reflection mapping of movies 2.

4.7.4 Film as Surface

Everything within the universe of *La Dispersion Du Fils* is constructed from moving images. This vague statement must now be fleshed out into an algorithmic account of how the video database is transformed into visible light.

The virtual optics described above prescribes exactly the various ways in which the moving images of the video database might be rendered visible. Specifically, video-data might a) be used to alter the properties of light (colour, intensity, direction) and be duly rendered visible when this light illuminates a surface, or b) the video-data might be used to define or alter the optical or geometric properties of a surface itself, and thus be revealed by perverting the passage of light.

The optical properties of a surface that may be altered or derived from video-data are diffuse reflectivity (albedo), specular colour and strength, emissiveness, bump strength and bump direction. Combining these can yield interesting results. For example, the brightness of the video-data can be used to modulate the strength of the bump-map, as illustrated in Figure 104.

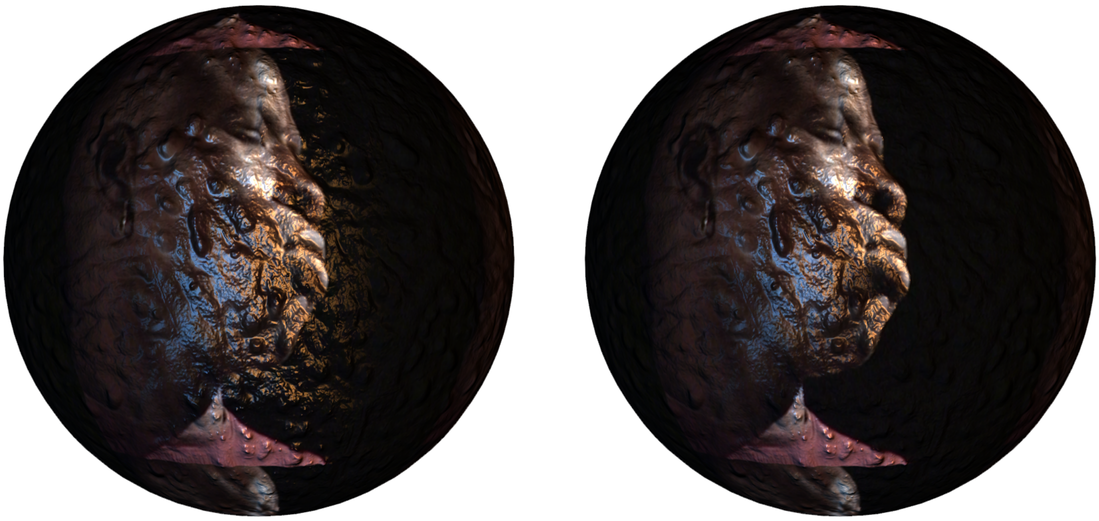


Figure 104 - Movie data modulates bump strength. Left: Movie-data is used as surface albedo only. Right: Movie-data is used as surface albedo and bump-strength.

4.7.5 Moving Images

In the manner described above, data contained within the video-database is used to define the various optical properties of the skin of the Helix. In this section, the mapping of this data to the surface of the Helix is discussed.

The video-database, when taken as pure numerical data, presents an inexhaustible source of organic patterns and structures. However, the information within these streams of bits that defines them *as moving images* lies predominately in the local spatiotemporal variation of the data. It is, then, the interpretations of these numerical streams that preserve local spatiotemporal structure - local changes in colour and brightness over space and time - that enable the video-data to be perceived as moving images. Therefore, although there are an innumerable variety of potential mappings of this raw numerical data to a two-dimensional surface, it is those that preserve this local structure and permit the perception of moving images that are sought.

One approach to ensuring the preservation of the local structure is to interpret each video as a 3D volume of pixels, with two dimensions being the image-dimensions (u and v) and the third dimension being time. Intersecting this 3D volume with any 2D surface will yield a 2D mapping in which a certain amount of spatiotemporal structure is preserved. One could, for example, construct a 3D grid of all these 3D-video-volumes, and then trace the intersection of the surface of the Helix through this volume of data. As the Helix undulated through this invisible volume of spatiotemporal data it would reveal the spatiotemporal images within, just as carved piece of wood reveals its rings.

Storing the video data in such 3D volumes, with random, instantaneous access to any set of spatiotemporal coordinates within, represents an ideal data-structure for *La Dispersion Du Fils*, for it leaves wide open a door for discovery of unexpected patterns and structures within the video data. It is also the simplest approach. There are, however, a number of technical constraints that make such an approach infeasible. The 526 films in the video database, when uncompressed into RGB images, require 400 gigabytes of storage (640 gigabytes when counting the panorama $CaMg[CO_3]_2$), far beyond the ½ gigabyte of VRAM or 4GB of RAM available on each render-node when this project began. Hardware texture compression (DXTC1) offers a 6:1 compression ratio, which goes little way towards reducing the size of the data-base to a level that might permit instantaneous random-access storage.

The solution adopted here is a hybrid system, where a portion of the video database is stored permanently in random-access format in VRAM, while the remainder is stored in a highly compressed state on hard-disk.

Helix Movies

378 clips, each exactly 3 seconds in length (75 frames), were selected from the database of films. Representative of the database as a whole, they were chosen for their

textural, motion and rhythm properties. The individual frames of each clip were extracted and combined into a single image mosaic, producing 378 mosaics, each with 5 columns of 15 frames. MIP-maps for each were computed offline, prior to mosaicing, to avoid any 'bleeding' of pixels between each frame during minimisation. Each of these mosaics were compressed using the DXTc1 compressed-texture format, and the entire process was repeated at a variety of resolutions ranging from 512x512 for a single frame (7680x2560 for a clip) down to 64x64 (960x320 for a clip). This permits the selection at run-time of an appropriate resolution given the VRAM capacity of the host computer.[208]

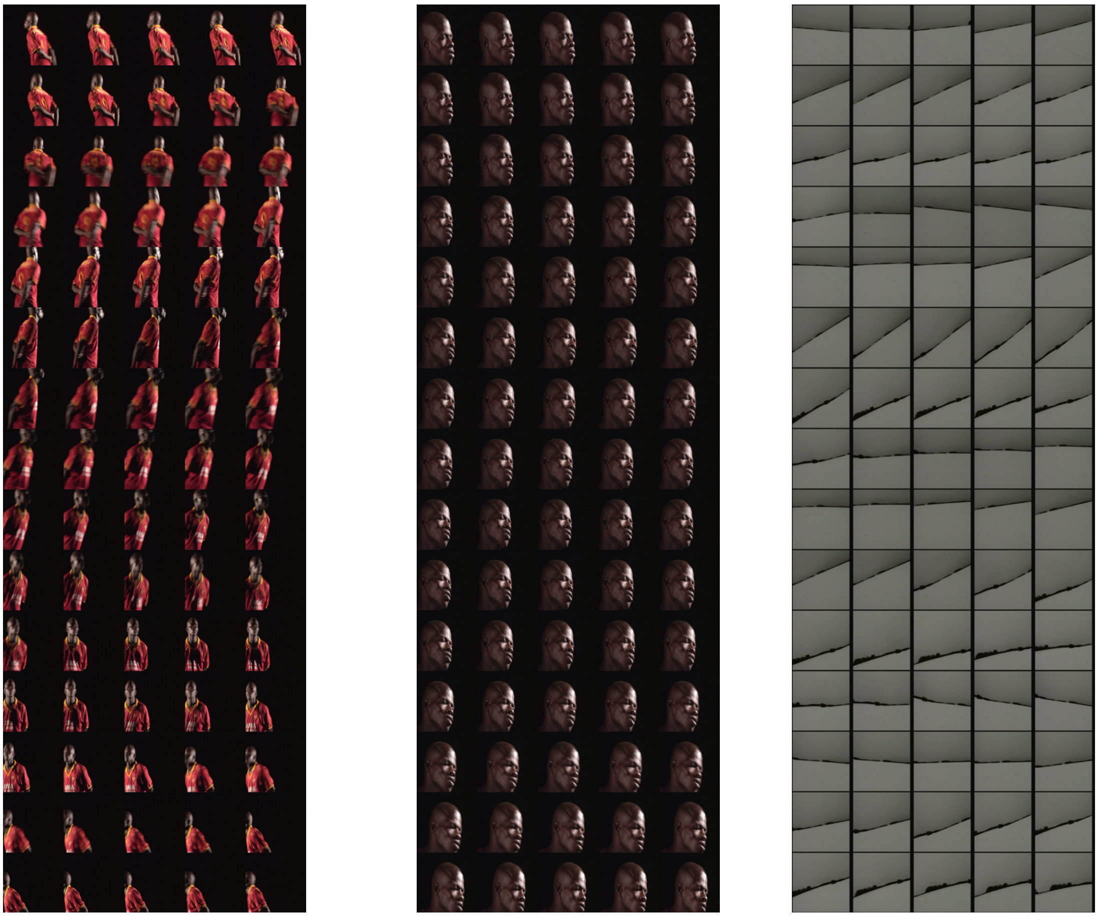


Figure 105 - Three Helix-movie motion-texture plates.

The result is a library of motion textures that can be instantly mapped to any 3D surface within the virtual world without hesitation. These 378 clips will be referred to as the *Helix-movies*, for they are chiefly destined to appear on the surface of the Helix.

[208] From 4.6GB for the highest resolution down to 73MB for the lowest.

MPEG Engine

In addition to the Helix Movies, a system for playing all of the movies in their entirety was developed. This body of work was not without significant challenges. However, to avoid any breaks in presence, it is critical that the flow of the Helix never be interrupted. A video engine capable of streaming large numbers of films onto arbitrary 3D surfaces without disrupting the overall frame-rate of the system was required.

To this end, a custom video engine was developed specifically for the multi-threaded streaming, decoding and display of numerous video streams on 3D surfaces without ever causing the host application to halt or stutter whenever new videos are launched, or the workload grows too high. The engine uses multiple threads to pre-emptively stream data from disk to RAM, bypassing the CPU. Decoding is performed on the CPU, again in dedicated threads. The copy from RAM to VRAM is performed only when the video frame is visible on the host computer. As each computer in the cluster sees only $1/6^{th}$ of the panorama, this copy-on-demand system is very effective in reducing the expense of moving data from RAM to VRAM. Once in the GPU, MIP-maps are calculated and finally, during rasterisation, colour-space conversion from YUV to RGB and further colour and gamma adjustments are performed.

The engine avoids all dynamic creation or destruction of memory structures, instead recycling a pool of RAM and VRAM frame-buffers. Most importantly, the frame-rates of the videos are uncoupled from the main application frame-rate. This allows rendering and presentation of the 3D scene to continue at 60 frames per second, regardless of the state of the videos. The frame-rates of each of the videos are also decoupled from one another. This has the advantage that, if the video-engine is ever saturated and unable to meet demand, it is typically just one or two videos that fall behind schedule, while the rest continue to play smoothly. Each video must, however, remain synchronised with itself across the cluster. The video-engine is, therefore, 'cluster-aware.' This allows videos to be displayed in synchronisation with soundtracks, should they have one, as sound is rendered on the master computer only.

MPEG2 compression was chosen for its computational efficiency and the availability of open-source implementations of the decompression algorithm. Most importantly, it was one of the few codecs with implementations that permit arbitrary resolutions, a necessary feature for encoding the panoramic video. In addition, being an older codec, MPEG2 favours computational efficiency over high compression. As CPU cycle prove to be the performance bottleneck rather than disk speed, this property of MPEG2 is highly beneficial.

The engine is capable of displaying approximately 10 - 12 HD video streams simultaneously. The videos in *La Dispersion Du Fils*, however, are compressed at much lower resolution, namely 352x288, allowing the system to stream up to 150 videos simultaneously, without stuttering.

The MPEG2 codec is, however, sequential, and in order to access one frame it is almost always necessary to decode the ones before it. A system for seeking specific video frames is provided, but the engine is by no means random-access, in that frames cannot be played in any order whatsoever. This demands of *La Dispersion Du Fils* that video data be used in a more or less traditional manner, with video frames displayed in their original temporal order. This constraint, however does not apply to the Helix-movies, where instantaneous access to all frames is available.

4.7.6 Montage by Motion

The data contained within images is used to define various optical properties of the Helix's skin, from colour to roughness or reflectance. In this section, the geometric arrangement of these images on the surface of the Helix is described.

As the viewer travels the length of the Helix, sections of the Helix are continually recycled from the tail to the head. During this stage of renewal, new images are selected from the database and mapped to the skin of the Helix. The films are mapped to the surface of the Helix in a grid-like pattern, following the 'quad' structure of the underlying polygonal mesh. Films may occupy more than one quad, but are always aligned along quad boundaries and never overlap. This way, every quad is covered by one and no more than one film.

The selection process works by forming 'probability-waves' of films. Each wave possesses parameters that dictate the density, the number and diversity of films, orientation, range of permissible sizes and temporal coherency of speed and phase. For example, one wave might suggest a choice of two different films, and that they be given the same orientation, equal physical size, and play in temporal phase, while another wave might offer a choice of 10 different films, be given random orientation size and exhibit no temporal coherency. These parameters are then triplicated, so that they may describe size, orientation, selection and timing of films differently at the beginning, middle and end of the wave. This allows for smooth transitions in the arrangement of films over the length of a single wave. For example, films at the beginning of the wave may have an initial size of just one quad, but then gradually grow along the length of the Helix.

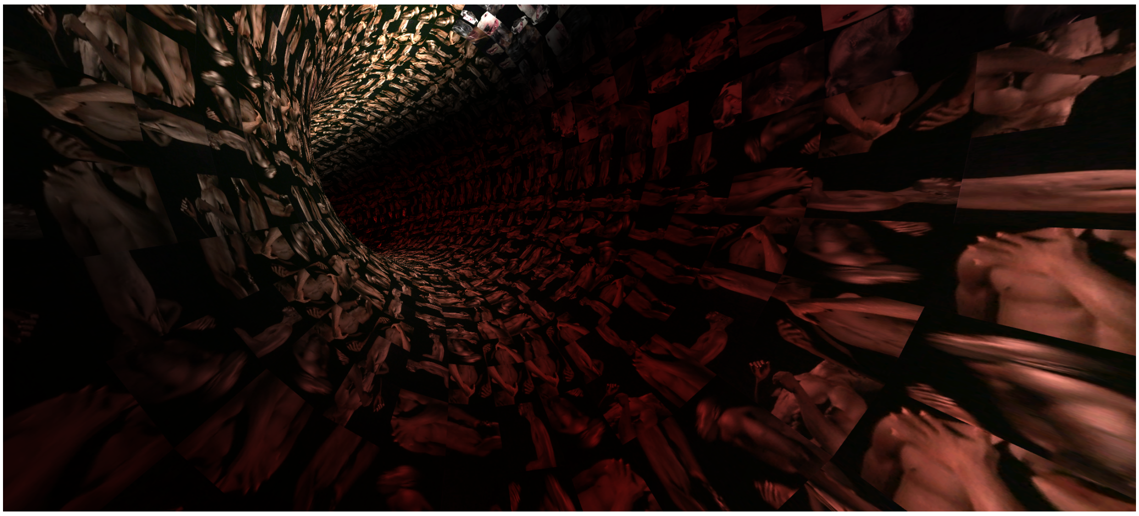


Figure 106 - Movies mapped to surface of Helix.

Numbers of these film-waves are superposed randomly along the spin of the Helix, such that all points along the spine lie under at least one wave, while some lie under two or more. Waves are given random lengths, ranging from many times the length of the visible Helix to as little as the length of one section (~60 metres). As the viewer advances along the Helix and new sections are regenerated, each polygon in the new section is assigned a new film by first selecting stochastically from one of the local film-waves according to the distribution of probability densities at this point on the spine. Once a choice between superposed waves is made, the wave's parameters are used to select a new film, size, orientation and temporal offset, by interpolating between the start, middle and end parameters of the wave. To accommodate films of arbitrary size, a form of 'box-filling' algorithm is used to ensure that no film is cut or cropped.

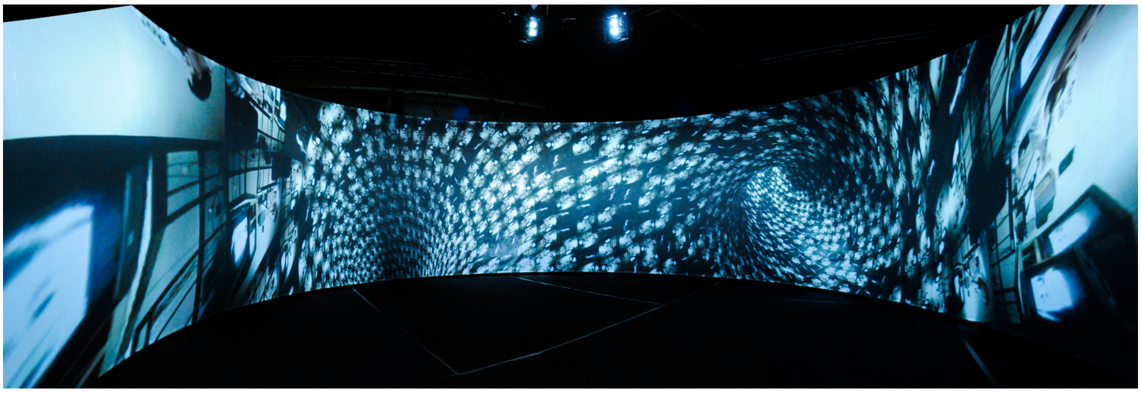


Figure 107 - Mapping movies to the Helix - emerging audiovisual patterns. Photo in AVIE.

In this manner, films are applied to the surface of the Helix. The result is a form of perpetually evolving tapestry of images, at times highly ordered, at others chaotic, but always in smooth transition as one wave of films gives way to another. The relative temporal synchronicity or asynchronicity of the films creates spatiotemporal rhythms, pulsations and rippling patterns across the surface of the Helix. These motions complement the already animated skin of the Helix, creating rich visual fields of motion texture.

When seen from a moving viewpoint, the effect is a form of dynamic montage, as the spatial relationships between the images are transformed into temporal ones. The viewer cannot resist but draw connections between the images as they wash upon them and narratives, in the loosest sense of the term, emerge. And with all the different patterns that the films can be arranged on the surface, and the hundreds of films to choose from, these haphazard journeys through the video database are destined to never quite repeat.

Later, it will be seen how this fabric of short, looping Helix-movies is augmented by periodic appearances of the movies in their entirety. Here, the MPEG2 engine is used to display 'windows' of video on the surface of the Helix, in which the films are shown without interruption. This is implemented using the same method of 'probability waves,' where some waves are given access to the full-length videos, thereby providing the viewer with access to the entire database of footage.

Figure 108 - Mapping movies to the Helix. Audiovisual rhythm and texture emerges from the haphazard spatiotemporal juxtaposition of films.

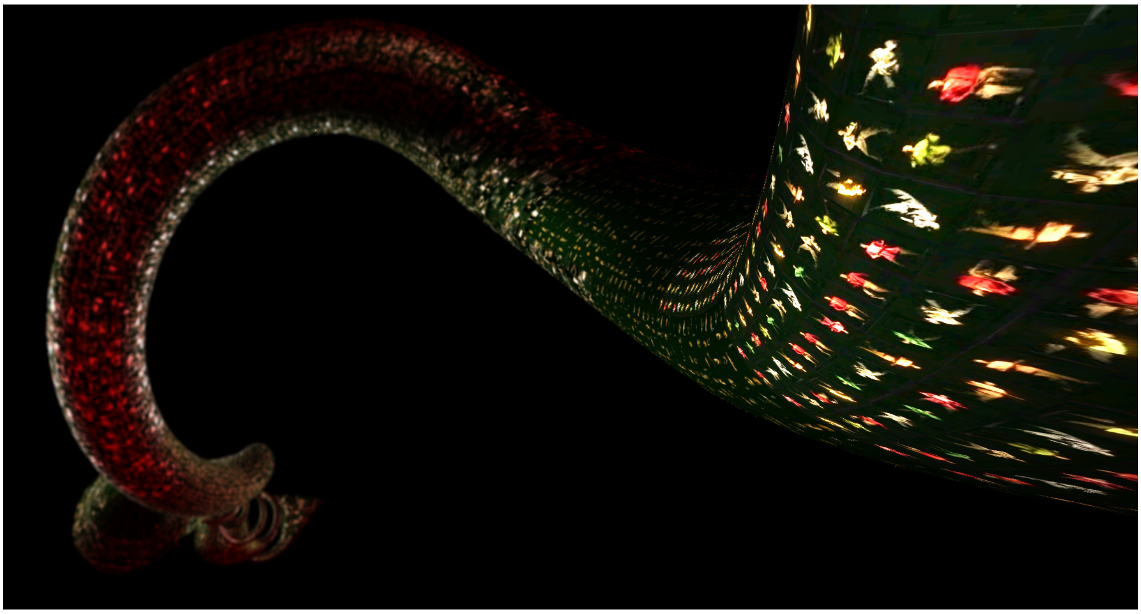


Figure 109 - Mapping movies to the Helix.

4.8 The Sound of the Helix

Each and every one of the 3-second Helix-movies were assigned a unique 3 second sound-clip. Hundreds of sounds were created by musician and sound artist Thierry Arredondo using his voice alone, and the result is a database of *sonic textures*, in many ways the exact sonic counterpart to the library of Helix-movies.

During the recycling of Helix section, as the Helix movies are mapped to the mesh, these Helix sounds are also attached to the surface of the Helix. As it is neither possible nor necessary to attach a sound to each and every Helix-movie, the ten largest

Helix-movies of each Helix section (or if they are equally sized, 10 randomly chosen Helix-movies) are given spatial sounds.

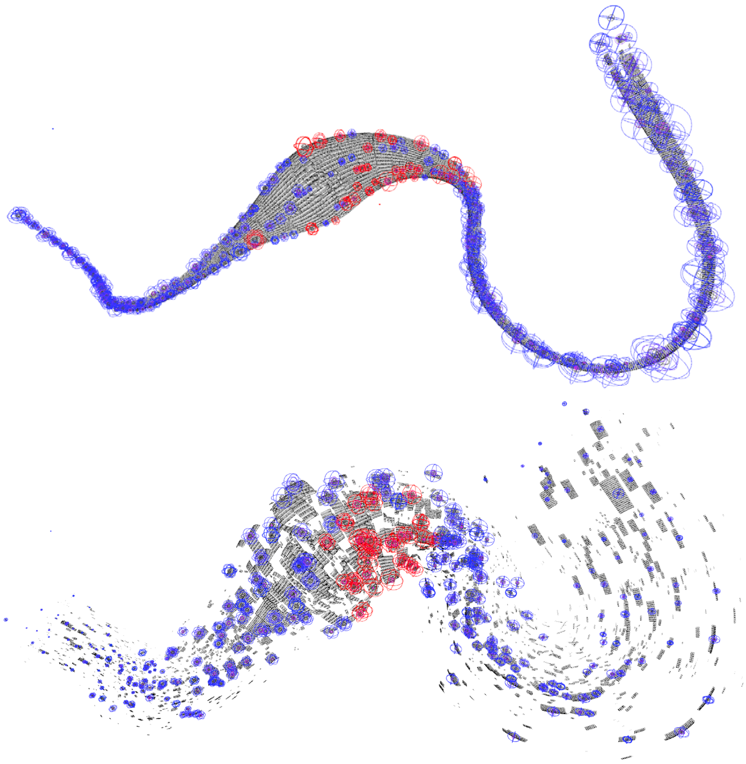


Figure 110 - Sounds mapped to surface of Helix. The spheres, both red and blue, are sounds attached to a subset of Helix-movies, with priority given to the largest Helix-movies. The sounds follow the skin-fragments as the Helix disperses or mutates.

Now as the viewer advances along the Helix, they travel through a spatial composition of sound textures that evolves over space and time in exactly the same manner as the mosaic Helix movies. The sounds are synchronised with the Helix-movies, so just as the temporal (a)synchronicities of the images creates spatiotemporal visual patterns, here these temporal relationships form cadences and syncopations, murmurs, throbs and echoes. As the Helix bulges, ripples, curls and distends, the sounds move with it, firmly attached to the inner surface of the skin.

The result is a form of algorithmic sonic composition created and experienced by ego-motion. When heard through the spatial sound system of the AVIE, the sound field has a tangible structure, and the sounds can be heard from afar and receding into the distance, giving the viewer an impression of their velocity. When the Helix curls in on itself, its structure can be heard through the walls of the Helix, as the future and past of the traveller's path pass within ear shot of the present.

In concert to this travelling train of sounds, a number of additional sounds are used to further enhance the experience of space and motion. A number samples evoking the sound of travel through an enclosed space are used, all of sufficient length to avoid any perception of repetition. These samples are pre-filtered with varying levels of reverberation and spectral attenuation, and as the Helix contorts or straightens, dilates or constricts, the relative levels of these sounds are altered to reflect the changing reverberant qualities of the Helix. When the Helix distends to form giant caverns, 400 metres in width, or shrinks to a single metre in width, these sounds reflect these changes. This method is extended by adding variations of the sounds that reflect different speeds of motion through the Helix, and these two are blended according to the current speed of the viewer.

A special sound is also reserved for the skin of the Helix, a pre-mixed cacophony of all the voices captured within the database. This sound is heard only when the viewer approaches closely the skin of the helix, peaking as they pass through it on their way in or out of the Helix.

The effectiveness of this sonic composition lies, above all, in the successful multi-modal binding of visual and sonic stimuli. It is this that allows the Helix to be perceived as physically present, or the viewer to feel physically present. It is in observing the immersive potency of this synaesthetic arrangement that a theoretical account of the role of binding in presence was given such importance in the first part of this thesis.

4.9 Dispersion of the Helix

The Helix exists in a constant state of growth and decay. At its head, fragments fly in from the darkness to assume their rightful position in the fabric of the Helix, where they stay until they are discarded, jettisoned into the ether at the Helix's tail.

However, this fragmentation is not limited to the head and tail. From time to time, the Helix may enter states of high fragmentation, in which the entire skin, or just the region around the viewer, splinters or shreds. This is implemented using a class of skin-modifiers called *HelixEffects* that operate on a per-vertex granularity, displacing the vertices away from the spine. Two variations exist, the first displacing the vertices along the vertex's normal, the other displacing the vertex along a shared face normal. The difference being that in the latter the size of each displaced face remains constant, while in the former the displaced faces grow with in size with displacement. The degree of displacement varies from Helix-movie to Helix-movie, z-coordinate along the spine (local or global) and with varying amounts of granularity. These displacements may be at times subject to animated noise, leading to organic motions, not unlike tentacles or seaweed in water.

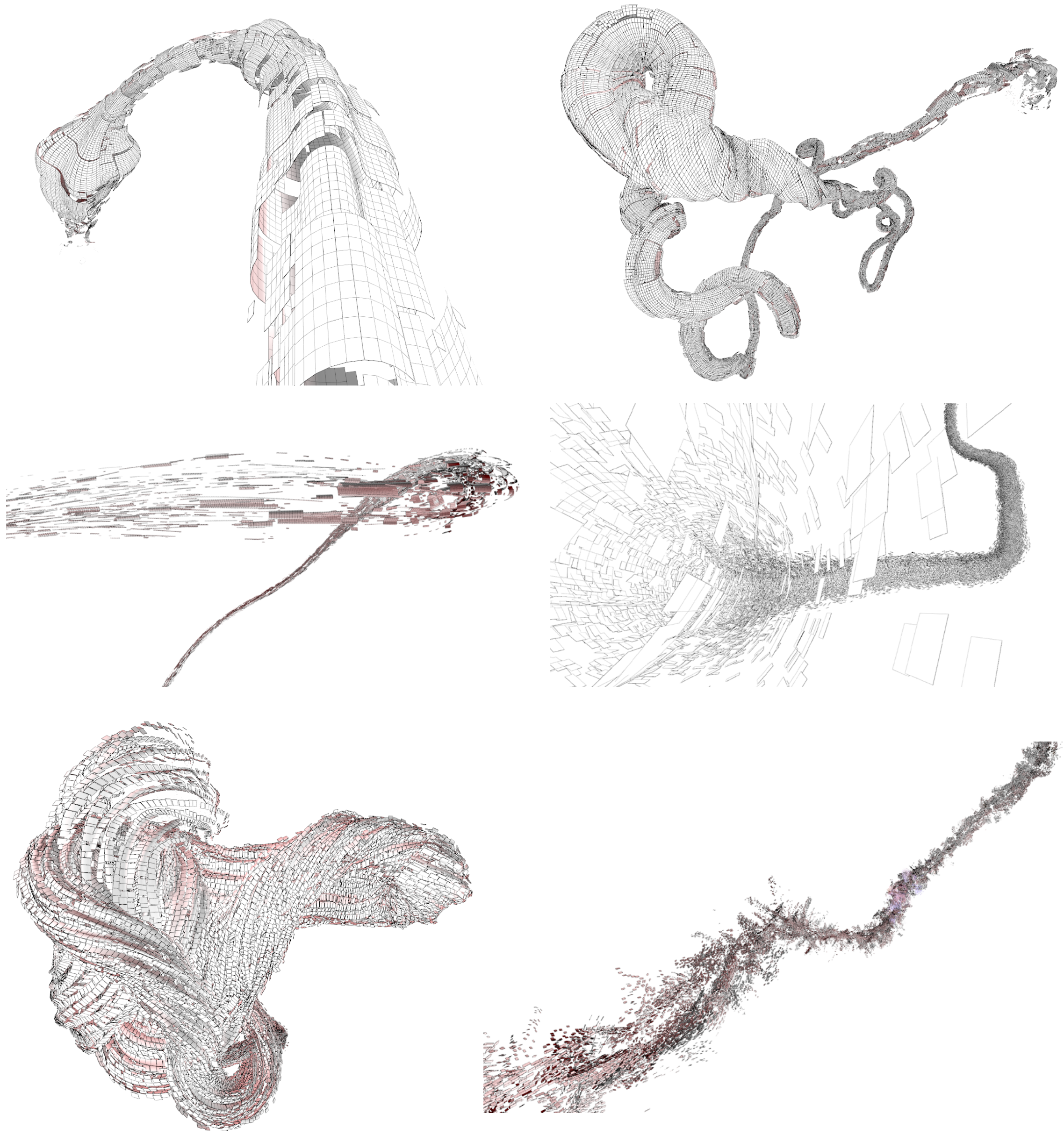


Figure 111 - Dispersion of the Helix.

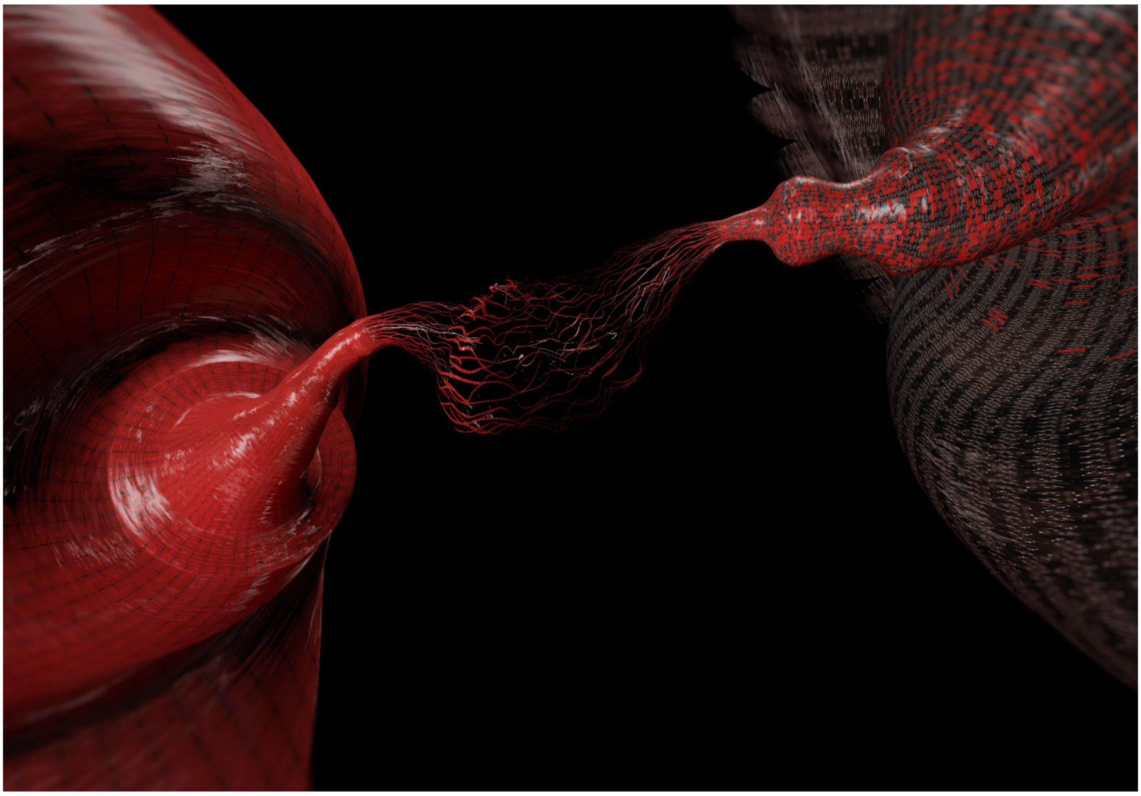


Figure 112 - Partial, string-like, dispersion of the Helix.

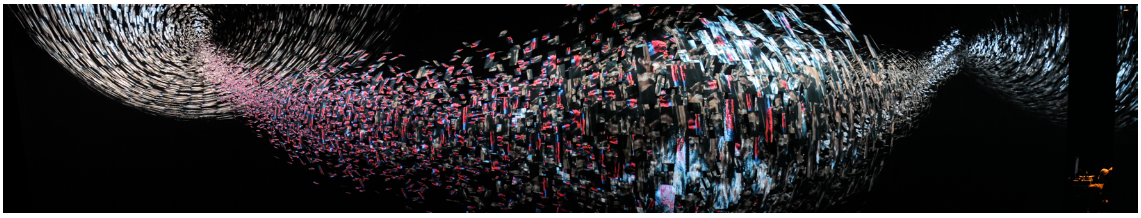


Figure 113 - Dispersed Helix. 360° panoramic photo captured in AVIE.

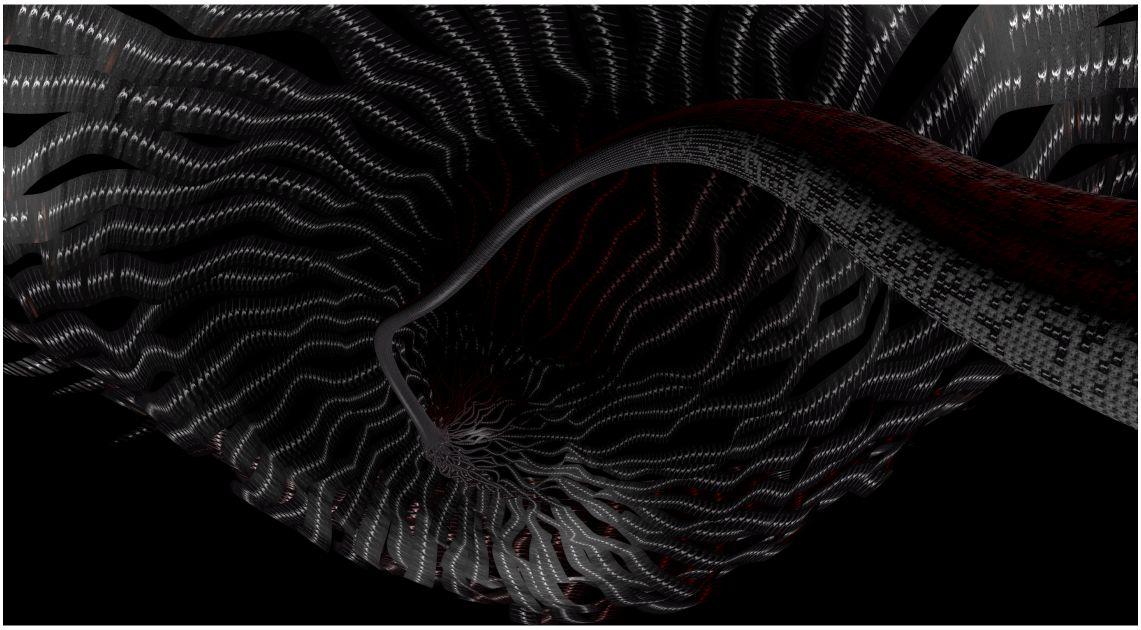


Figure 114 - Ribbon like fraying of the Helix combined with the Flare modifier.

4.10 EGO-MOTION

The essential experience of *La Dispersion Du Fils* is a journey through, and with, the Helix. To achieve this motion, the AVIE is modelled as a moving platform, with mass, stability and velocity. The AVIE is bound to follow the Helix, but the character of its motion is constantly varying. Sometimes rapid, other times slow. Sometimes tumbling, or pitching and tilting, or at other times its orientation fixed in the direction of travel. The large part of the viewer journey takes place within the Helix, but forays into the outside world occurs sporadically. Occasionally the AVIE will drift a great distance off into the void, and a distant perspective of the Helix is won.

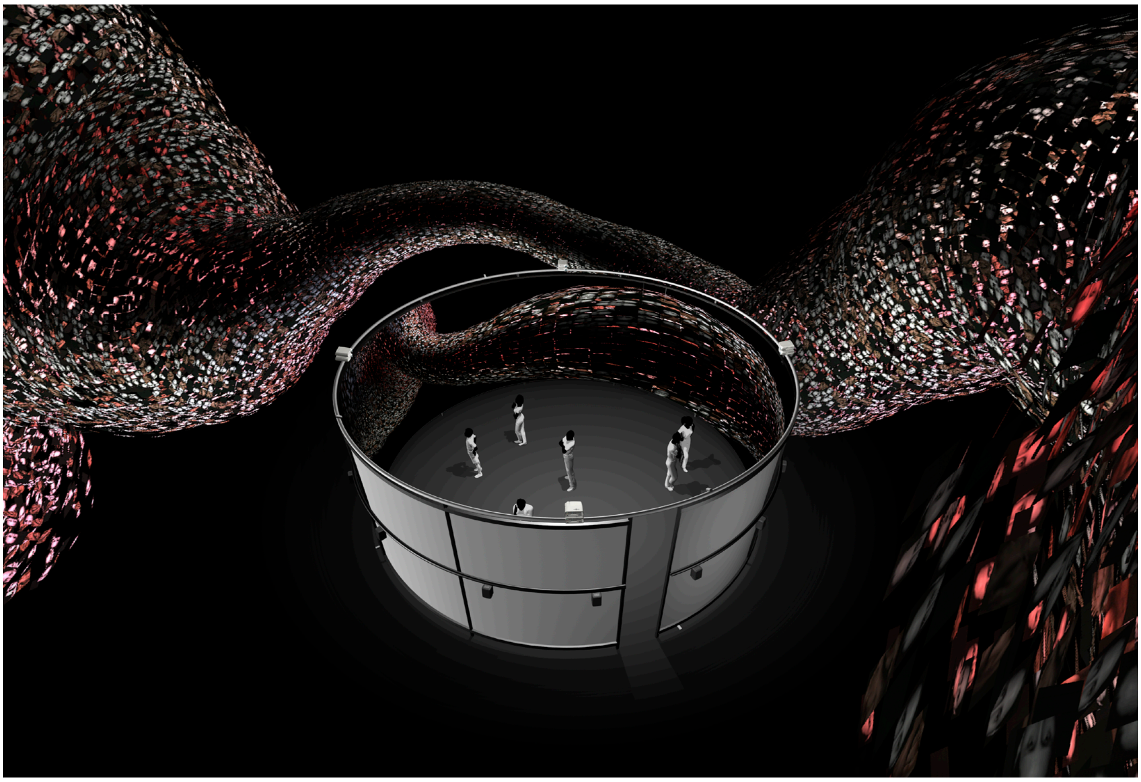


Figure 115 - The AVIE as vessel.

In many respects, the whole exercise can be considered an experiment in vection and the perception of rest-frames: continuous manipulation of speed, size and frequency and rhythm of motion stimuli, linear motions and circular motions, from within the Helix where the entire field-of-view is engulfed in optic flow, or from without, where a variety of potential rest-frames compete for purchase. When the Helix fragments, dense fields of motion parallax cues and foreground-background relationships are presented. Added to this, a coherent moving spatial sound field, allowing for perception and binding of multi-modal vection cues.

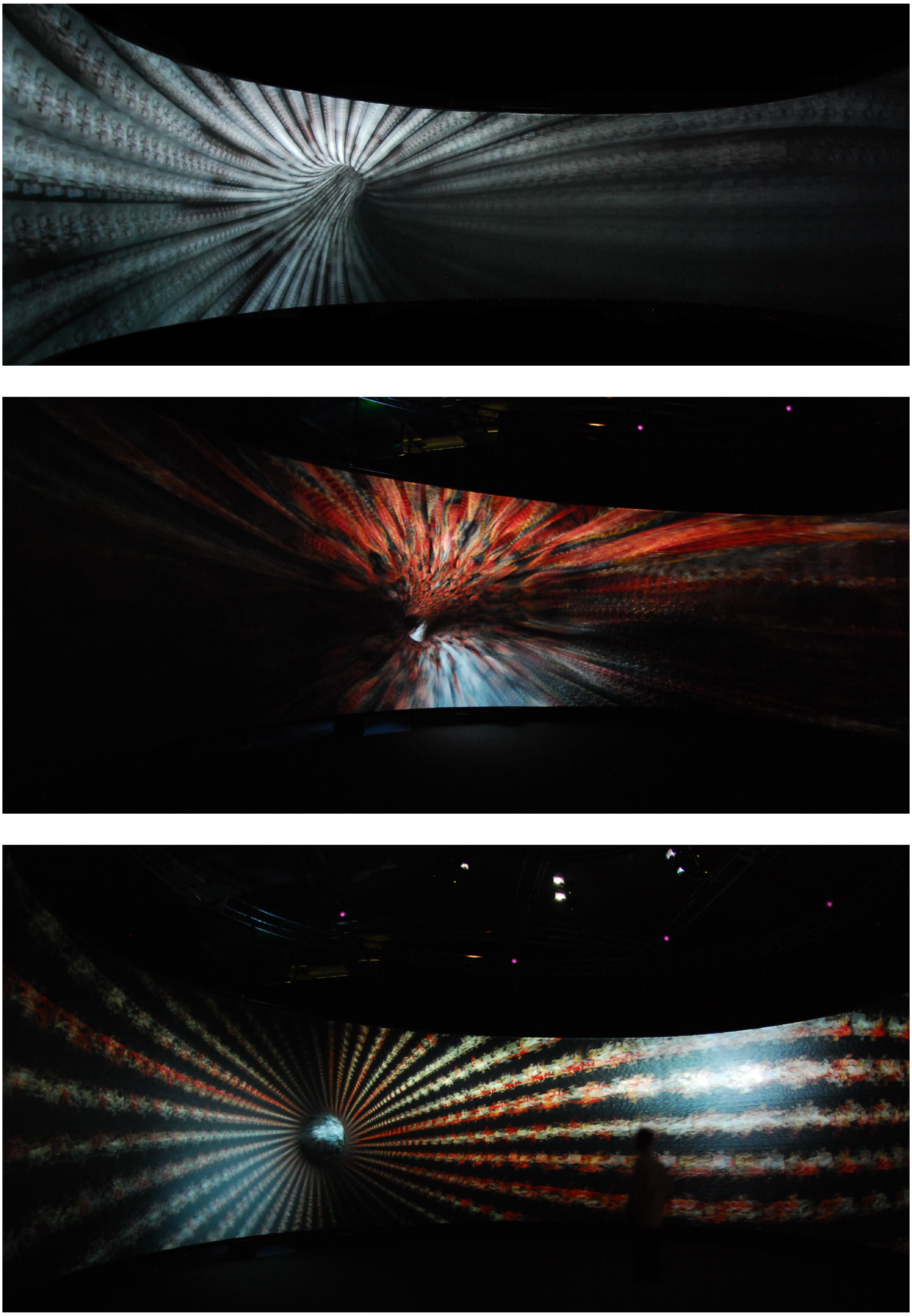


Figure 116 - Helix motion-textures. Long-exposure photographs captured in AVIE.

4.11 MUTATION AND METAMORPHOSIS

The state of the Helix is determined by the combined influence of many dozens of parameters, controlling such things as its appearance, motion, growth and speed, movie textures, sound, and so on. In fact, every aspect of the Helix is parametric, and all of these parameters are open to interactive manipulation. This deliberate feature of the software was originally intended to provide the artists with a means of interactively 'designing' the Helix, at times from within the AVIE itself, searching for the 'best' combination of parameters to be set in stone.

As development advanced, however, it became apparent that it was far more interesting to leave the setting of these parameters open to chance. And very soon this became a policy, that *none* of the parameters describing the Helix, the motion of the AVIE, the appearance of phenomena or Circles, should be pre-determined or fixed. Everything that could be, would be opened to randomness.

To effect this, a library of *HelixMutators* was introduced, autonomous objects that determined how other parameters change over time. For example, the *NavigationMutator* is responsible for altering the parameters that control the motion of the AVIE. The *HelixSpineMutator* alters the parameters that determine the path and movement of the spine, and so on. The parameters under the control of a mutator can be considered as a single multidimensional state-vector, which helps to explain the two ways that mutators generally go about their function. In one mode of operation, at infrequent intervals, one or a small number of dimensions of the state-vector are assigned random target values and over a period of time, the state-vector is interpolated towards this target state. Once this state is achieved, a new random coordinate is state-space is selected, and interpolation begins anew. In the second mode of operation, a random walk through state-space is performed, where at each time step, the state-vector is nudged in a direction determined by an animated noise-field. The significant difference between the two is that in the first mode of mutation, only one or a small number of dimensions are in motion at any time, such that the path through state-space is largely composed of straight lines, parallel to the axes of the space. In the second mode, all dimensions of state are almost always in motion, and the trajectory through state-space here is more like that of the Helix through 3D space – a sinuous, meandering path.

Figure 117 - The Helix in one of its innumerable forms.

The result is a world of permanent metamorphosis. The state space has a sufficiently high dimensionality to guarantee that no particular state will ever be visited twice, and that the same path through this space will ever be taken twice. The challenge is in finding parameterisations of the world that, when the parameters are altered, give rise to continuous, well-behaved change in the visible world. Further, the world must be well-behaved for all regions of state-space. These, I suggest, are fundamental challenges facing any algorithmic approach to artistic discovery. Unfortunately there is insufficient space to pursue this discussion further here.

4.12 THE CIRCLES

During October and November of 2008 *La Dispersion Du Fils* was presented publically for the first time, as part of the Shanghai eArts Festival. This premiere version of the work was marked by two distinguishing aspects that would not feature in later versions: *interactivity* and a series of phenomena known as *the Circles*.

In the Shanghai version of the work, the Helix serves as a conduit between phenomena known collectively as *the Circles*. The viewer travels along the Helix, emerging sporadically into one of 31 unique Circles; large arenas and spatial structures within which the viewer is given an opportunity to watch the films in their entirety.

The structure of these Circles, and the Helix providing passage between them, echoes the earliest of human constructions: megalithic stone circles and cromlechs, henge and tumuli, barrows and dolmen. These Neolithic structures, however, are never graphically depicted or imitated. Rather, it is the arrangement of stones - from the small haphazardly packed stones of the narrow sunken passages to the giant monolithic slabs of the stone circles - that inspires the distribution of films across the skin of the Helix and within the Circles.

Like their megalithic inspirations, the Circles are sacred places, funerary places, places of remembrance, of death and transcendence. In *La Dispersion Du Fils*, the Circles serve as portals into the worlds captured within the films of LFKs. The 526 films are divided among the 31 Circles according to their thematic, narrative and aesthetic concerns.[209] Most importantly, and in contrast to the 3-second looping clips used to construct the Helix, the films in the Circles are shown in their entirety. As the viewer travels from Circle to Circle, they pass through a spatial-montage that, together with the Helix, constitutes the 'grand narrative' of the work.

The Circle films are displayed on the surface of blocks, the dimensions of which vary greatly across and within the different Circles. Often the blocks are scaled to form rectangular slates or slabs, ranging from one or two metres in size to tens of metres. In

[209] The number of films in each Circle varies greatly, with four of them possessing just one film, while others possess as many as 76 unique films.

some cases the film surfaces are treated as sources of light in themselves and so do not require illumination from other sources of light in order to be seen, while in others, the film-data is treated as the diffuse albedo of the surface, and made visible by scattering surrounding light.

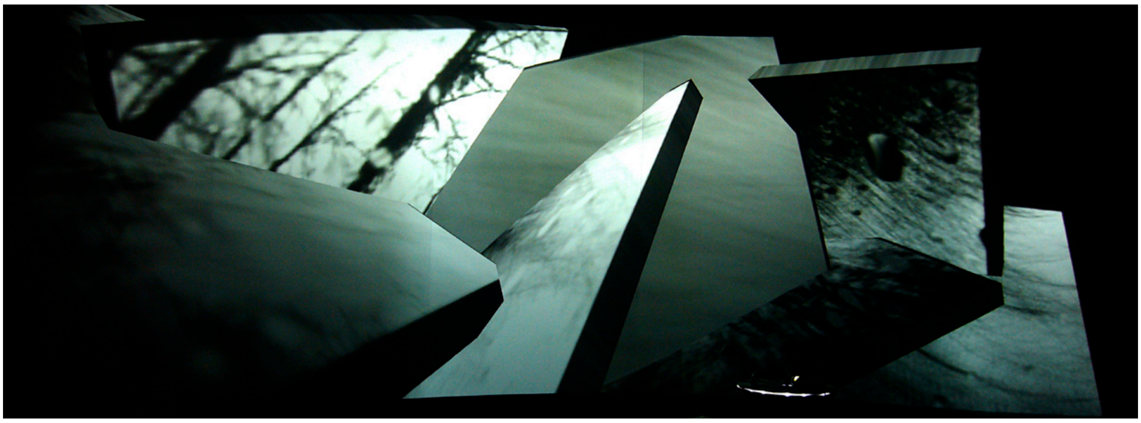


Figure 118 - Circle VII.1 Elements – Forets. Photo taken within the AVIE.

The films are put in motion around the viewer and are never stationary. The movement and arrangement of the films is algorithmic, and the algorithm governing the motion of the films is unique to each Circle. These algorithms were hand-crafted in accord with the thematic concerns of the films themselves. For some circles, these algorithms include simple simulation of forces such as gravity or turbulence, or attraction or repulsion between films themselves, while in others, the movements are not imitative of any obvious natural phenomena. In all cases the algorithms possess a certain element of randomness and noise, ensuring that the motion and position of the films never repeats.

The design of all 31 Circles cannot be described here, but it will suffice to provide a handful of examples. In Circle III.1 *7 Chiens,* the seven dogs are replicated innumerably, and presented as torrent of images streaming towards and around the viewer, as does a river around a rock. In VII.7 *Elements - Dispersion*, the films violently and repetitively lunge and thrust at the viewer from all directions.

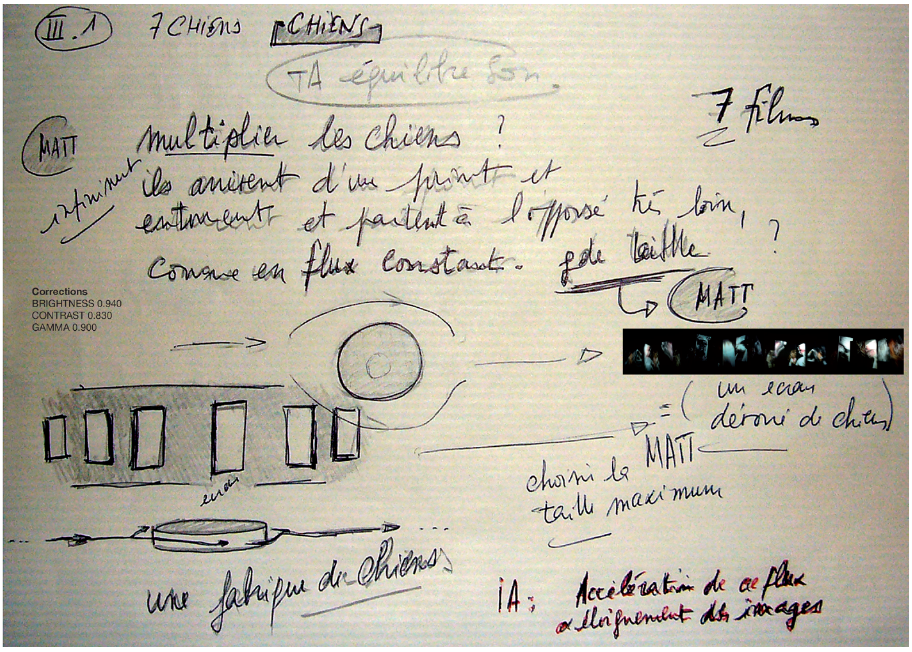


Figure 119 - Circle III.1 7 Chiens. Sketch by Jean Michel Bruyère.

In IX.2 *Enfants de Nuit - Portraits*, the films float down from the above, tumbling and swaying gently as if falling through fluid, while in X.6 *L'Insulte Faite Au Paysage - Courses* they descend in a rapid cascade of images.

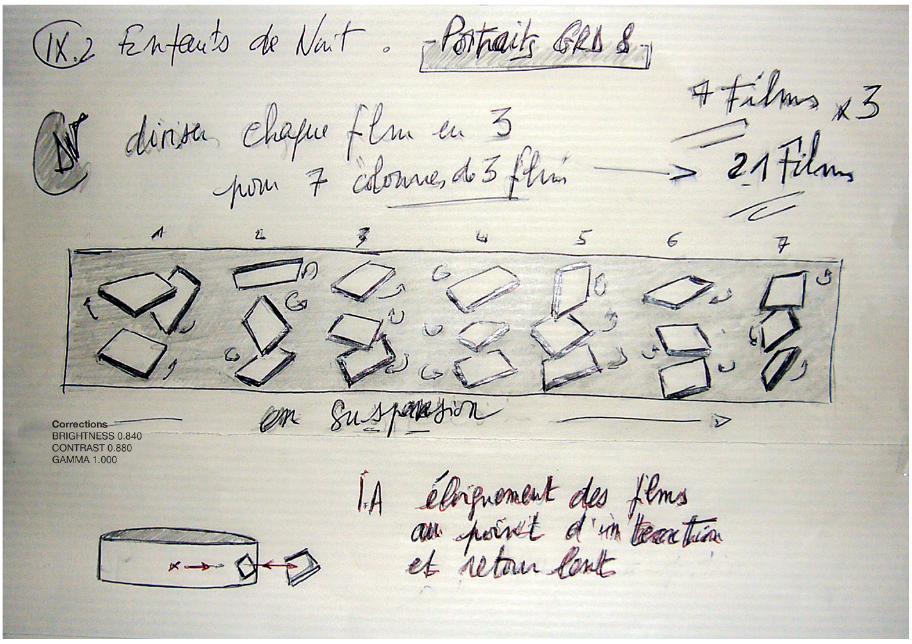


Figure 120 - Circle IX.2 Enfants de Nuit. Sketch by Jean Michel Bruyère.

In VII.1 *Elements–Foret* the films jostle and intersect one another, while in VII.6 *Elements–Issa,* the films are mapped on to the sides of immense cubes, which slowly tumble around various axes, as they rotate around the viewer.

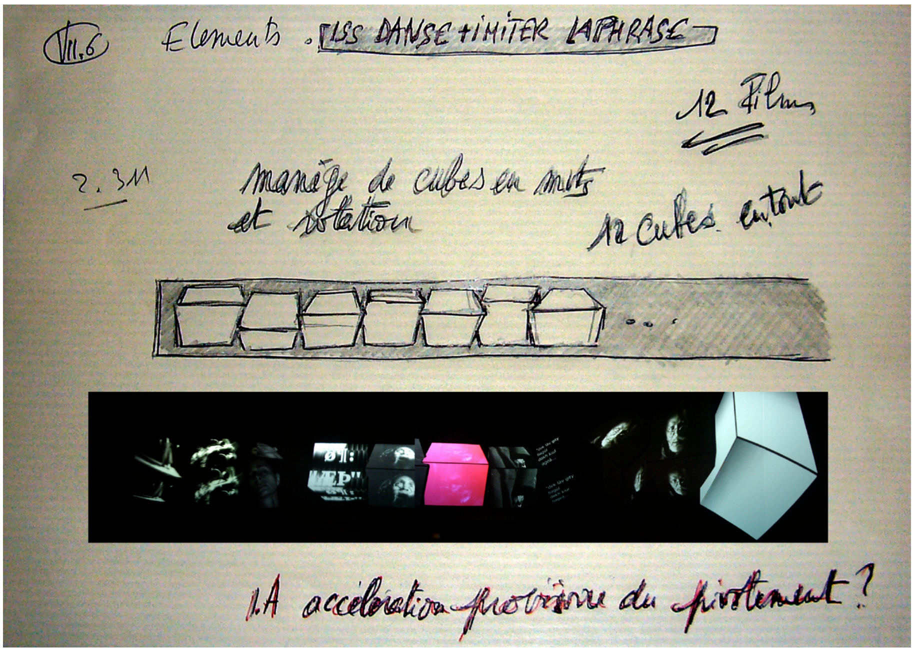


Figure 121 - Circle VII.6 Elements – Issa. Sketch by Jean Michel Bruyère, with panoramic photo taken in AVIE.

Circular Sound

Many of the circle films possess their own unique soundtrack and these soundtracks are attached spatially to the films, so that as the film-objects move through space, their soundtracks move with them. In this manner, the Circles are as much, if not more, aural experiences as they are visual, for the rich soundtracks of the films, some with dialogue or music, are physically arranged into moving, spatial sonic 'sculptures.'

Film as Object and the Immediacy of Space

In his essay Cinema and the Code (1989), Gene Youngblood is discussing the impact (then and future) of the computer on the moving image when he suggests that the formal possibilities offered by digital images "as syntactic elements or linguistic primitives" are "1) image transformation, 2) parallel event-streams, 3) temporal perspective and 4) the image as object." It is the last of these possibilities - the image perceived as object - that captures the essence of the Circles. However, it is the *duality* of vision afforded by the image-object that provides the Circles with their intrigue. The Circles, in inviting the simultaneous perception of both the images and the image-objects, present the viewer with a multiplicity of visions: those captured within the images themselves and the physical space and structure of the image-objects. This double vision brings the difference between perception of images and perception of space into stark relief. For one, the two-dimensionality – the superficial and artificial nature - of the images themselves is made all the more evident. But it is more than this; the image-objects are perceived as *immediately present*, as occupying space and possessing mass, momentum and substance. (Here, the spatial sound helps enormously to solidify their materiality and physicality). They are very much the *here* and *now*. In contrast, the scenes depicted within the images drawn on the surface of the image-objects appear remote and removed. They can *only* be seen as mediations or representations of distant or long-passed events. They have lost their *immediacy*.

In this way, the Circles provide space for reflection on the difference between mediated and immediate perception. They also suggest an important role for the traditional two-dimensional image in virtual reality: images, when presented as objects within a 3D scene, provide an aesthetic device for inducing reflective and critical distance. In an immersive medium, images allow the depiction of events outside of the here and now.

4.13 INTERACTIVITY

Interactivity, along with spatial audio, stereoscopy and real-time imagery, is one of the properties of the AVIE that distinguishes it from its predecessors. It is also, in addition to the Circles, the second defining feature of the Shanghai edition of *La Dispersion Du Fils*.

In the first part of this thesis it was reasoned that interactivity may, under certain conditions, increase immersion. More specifically, it was argued that interaction, in so far as it aids active perception of the environment, may enhance immersion. It was, however, also suggested that interaction could just as easily destroy immersion.

This section recounts experiments in interactivity with *La Dispersion Du Fils*. Using a custom-built motion-tracking system, the viewers are given some measure of influence over the unfolding virtual events. This is achieved by coupling some measurable property of the real world (i.e. the viewers) to some variable aspect of the virtual world.

Working with the premise that the most successful interaction strategy may not be the most obvious, a method for rapidly searching the space of all possible real to virtual couplings was constructed.

4.13.1 Monocular Tracking System

As Sridhar's *Immersitrack* tracking system (Section 3.12.1) was at the time in development, a simple single-camera vision-based tracking system was created specifically for the needs of the project. Created by fellow iCinema engineer Ardrian Hardjono, the system employed a single wide-angle camera suspended above the AVIE to obtain a birds-eye view of the panoramic arena. The camera operated in the near infra-red, allowing a number of infra-red lights to ensure adequate and constant illumination imperceptible to the human eye. Using the OpenCV blob-tracking modules with an interface written in Qt (C++), Hardjono created custom software for the real-time tracking of the two-dimensional position of audience members within the AVIE cylinder.

With this system in place, a single challenge arises: how can the position of the audience be used to influence change and trigger events within *La Dispersion Du Fils?*

One strategy involves the direct coupling of some measurable property of the real-world with some variable parameter of the virtual world. For example, the speed with which the virtual viewpoint advances along the Helix could be dictated by the physical distance of a viewer from the centre of the AVIE. To understand the potential of this strategy, it is necessary to identify all the measurable properties of the real-world, as well as all the variable parameters of *La Dispersion Du Fils* amenable to audience control. We begin with the former.

4.13.2 Interactive Mappings

The tracking system reports the position of each viewer within the AVIE. The information carried within this raw data is made evident by constructing two tables:

Instantaneous Trackable Quantity	
`Position`	AVIE-centric position (2d)
`Radius`	Radial distance to centre
`ScreenDist`	Distance to AVIE screen
`ScreenPoint`	Closest point on AVIE screen (2d)
$\texttt{Proximty}_{\texttt{MIN}}$ $\texttt{Proximty}_{\texttt{MAX}}$ $\texttt{Proximty}_{\texttt{AVG}}$	Closest, furthest and average distance to another audience member
`Time`	Time since entering AVIE
`HxSpine`	Helix-centric position (3d: z, r, θ)[210]
`HxSkinDist`	Distance to Helix skin
`HxSkinPoint`	Closest point on Helix skin (3d)

Table 5 - Instantaneous trackable quantities.

Tracking Operators	
`min(S, Q)`	Lowest Q value for all people in set S
`max(S, Q)`	Highest Q value in S
`median(S, Q)`	Median Q in S
`sum(S, Q)`	Sum of all Q across S
$\texttt{viewer}_{\texttt{MIN}}\texttt{(S,Q)}$	Select viewer with lowest value Q in S
$\texttt{viewer}_{\texttt{MAX}}\texttt{(S,Q)}$	Select viewer with highest value Q in S
`threshold(min, max,S,Q)`	Select viewers from S with Q between min and max
`count(S)`	# people in a set
`circle(S)` `ellipse(S)` `line(S)`	Circle, ellipse or line of best fit.
`delta(Q)`	Change since last reading
`ddt(Q)`	Rate of change
`d2dt(Q)`	Rate of rate of change
`magnitude(Q)`	Absolute magnitude
`accumulate(t,Q)`	Sliding accumulated value over time t
`average(t,Q)`	Sliding average value over time t

Table 6 - Tracking operators. `S` *is a set of people. If not specified, then all people in AVIE.* `Q` *is a quantity from Table 5.*

In the first table we see the instantaneous measurable properties of each member of the audience, namely the distance between viewers and various features of the real-world (the AVIE screen, the centre, other viewers) and the Helix (its spine and skin) as well as the length of time passed since the viewer entered the AVIE.

The second table provides a list of operations for aggregating, averaging, selecting, counting, differentiating or integrating a value from the first table, over multiple viewers or over time. For example, the average time spent in the AVIE of all participants can be written as `median(-, Time)`. (Here, the '-' indicates the set of all people within the AVIE, and can be omitted for brevity). The closest distance between all viewers and the centre is `min(Radius)`, or the sum of distances between all viewers and the skin of the Helix: `sum(HxSkinDist)`. The average distance between people,

[210] The position of the audience member in Helix-space.

which gives a measure of the *dispersion* of viewers with the AVIE, is given by $\texttt{median(Proximity}_{\texttt{AVG}}\texttt{)}$.

The `delta` operator returns the change in a value since the last reading while the `ddt` and `d2dt` operators provide rate of change over time. When applied to `Position`, they describe the velocity and acceleration of a viewer within the AVIE. The `magnitude` operator returns the absolute magnitude (i.e. vector length) of a value. For example, the scalar speed of a viewer is the familiar `magnitude(ddt(Position))`.[211] The most commonly used combinations of operation and property are provided with their own names for brevity: `Velocity = ddt(Position)`, `Acceleration = d2dt(Position)` and `Speed = magnitude(Velocity)`. Note however that the `ddt` and `d2dt` can be used with any quantity, and not just the position of the viewer. For example $\texttt{ddt(max(Prox}_{\texttt{MAX}}\texttt{))}$ gives the rate of change of the largest distance between any two viewers.

The `accumulate` and `average` operators provide the sum and average of a value over a specific time period. For example, the total distance covered by a viewer while in the AVIE is given by `accumulate(delta(Position), ∞)`, or the average speed of the group as a whole over the last x seconds by `average(median(Speed), x)`.

The $\texttt{viewer}_{\texttt{MIN}}\texttt{(Q)}$ and $\texttt{viewer}_{\texttt{MAX}}\texttt{(Q)}$ operators return the viewer with the lowest or highest value of a quantity `Q`.[212] For example, the viewer who has been in the AVIE the least amount of time is returned by $\texttt{viewer}_{\texttt{MIN}}\texttt{(Time)}$. These two can be used to chain operators together: the speed of the viewer who has been in the AVIE the least amount of time is given by $\texttt{Speed(viewer}_{\texttt{MIN}}\texttt{(Time))}$, or the position of the most isolated viewer is $\texttt{Position(viewer}_{\texttt{MAX}}\texttt{(Proximty}_{\texttt{MIN}}\texttt{))}$.

The `threshold` operation returns the set of viewers who fall within some range or threshold for a particular property. Those viewers who have been in the AVIE for more than 60 seconds, for example, is written as `threshold(60, ∞, Time)`. The

[211] `Radius` is simply shorthand for `magnitude(Position)`.

[212] And `min(S,Q)` and `max(S,Q)` are shorthand for $\texttt{Q(Viewer}_{\texttt{MIN}}\texttt{(S,Q))}$ and $\texttt{Q(Viewer}_{\texttt{MAX}}\texttt{(S,Q))}$.

`threshold` operator therefore provides a simple means for sub-dividing the AVIE into sub-regions. For example, `threshold(0, 1, Radius)` selects only those viewers within a 1-metre radius circle at the centre of the AVIE. Their aggregate velocity, for example, would then be given by `sum(threshold(0,1,Radius), Velocity)`. With this one could divide the AVIE into a number of regions, each with a different interactive mapping. One region might map the velocity of viewers to virtual speed, while another might map average position to virtual direction, for example. By visibly outlining such regions on the floor of the AVIE arena, and providing instructions to the audience, such 'control-zones' can provide individual viewers with finer interactive control.

Finally, the `circle`, `ellipse` and `line` operators perform a geometrical 'best fit' of their eponymous shapes on the positions of a set of viewers.

Real-world Events

A slightly different strategy for providing the audience some control over the unfolding narrative involves the detection of events within the real world, which can then be used to trigger events or episodes in the virtual world. Detectable events in real world include audience members entering or exiting the AVIE, entering or exiting some sub-region of the AVIE arena, touching or separating from one another, touching the AVIE screen, coming to stand still, starting to move, or breaching some threshold in speed, velocity etc. Such events can be reduced to the detection of changes in the verity of equalities written with this simple tracking schema. Someone entering or exiting the AVIE, for example, is expressed as `Delta(NumPeople) != 0`. The kinetic energy of the audience remaining above a particular threshold (`e`) for more than 10 seconds is `average(10, median(Speed)) > e`. An event is triggered when this statement changes from false to true.

Virtual-World Parameters and Events

Any parameter of the software that can be externally manipulated in real-time represents a candidate for interactive coupling. As the algorithms employed in *La Dispersion Du Fils* were deliberately designed to be highly parametric, almost any visual or sonic aspect of the system can be interactively driven. In fact, the number of parameters is too great to warrant an exhaustive listing here. Instead, a brief overview will suffice, in which the virtual parameters are divided between three categories.

In the first category, viewers are given control over some aspect of their flight through the virtual world, such as the speed of passage along the spine of the Helix, or distance from the spine or the skin, or the pitch, roll or yaw of the AVIE. For example, the average position of the audience can be used to tilt and bank the AVIE, as if the AVIE were a small water-borne vessel. This pitching and rolling of the AVIE can then influence the direction of travel, allowing the audience as a whole to steer a path through the virtual world.

In the second category, the audience's behaviour is linked to some variable property of the Helix itself, such as size, shape or curvature of spine, texture or lighting of the skin, speed and character of movement, level of dispersion or rate of mutation. For example, the kinetic energy within the AVIE (the sum of velocities of all audience members) could be mapped to some aggregate control of agitation or aggressiveness of the Helix.

In the last category, the audience's actions are linked to any of the large number of parameters that control the motion and display of film-objects within the Circles.

Note that not all parameters will produce instantaneous and visible change when altered. There are many parameters that control the rate of change of other parameters, for example, and changing these will only produce visible results after time. For the Helix, there are many parameters that only influence the growth of new sections at the head of the Helix, which will often not be seen until the viewers travel further along the Helix.

Events in the virtual world, in contrast to the controllable parameters, are relatively few, for change in *La Dispersion Du Fils,* while perpetual, is seldom sufficiently abrupt that it might be called an event. What events there are include the appearance or disappearance of the Circles, sudden motions of the camera, or sudden acts of mutation of the Helix.

Searching the space of all interactions

The reason for introducing this modular tracking schema is that it allows for *rapid* and *exhaustive* exploration of all the different ways the real-world might influence unfolding events in the virtual world. This approach is adopted on the premise that the most effective interactive strategies may not necessarily be obvious and may only be identified through trial and error. This framework clearly delineates the space of all possible interaction strategies as a mapping between the measurable quantities and events in the real-world and the properties and events of *La Dispersion Du Fils* amenable to interactive control. A search for effective interactive strategies is now a matter of *searching this space of mappings*.

To conduct a search, however, one must have some notion of what it is one seeks. What makes any particular interactive mapping effective or ineffective? What is the purpose of interaction? Why introduce it at all? In the context of this thesis there are two reasons. To begin, an understanding of the interactive potential and limitations of the AVIE system is sought. In particular, how the compromises and imperfections of the system impact upon, and are affected by, different interaction strategies. The AVIE also presents an opportunity to study the interplay between immersion and interaction in general, and an opportunity to put the theories proposed in Section 2.5.7 above to the test. In this regard, the goal is to seek mappings that directly impact presence, positively or negatively, and to understand why. In particular, the AVIE presents unique space for the study of multi-user interaction strategies, in which the actions of many are aggregated to influence a shared experience. With the operators that aggregate or select properties of a set of viewers serving as the basic building blocks of a multi-user interactive mapping, the tracking schema presented above is designed to allow the rapid construction of group interfaces in which the group as a whole is treated as a single interactive agent.

In the context of the artwork *La Dispersion Du Fils*, interaction is employed as an aesthetic device. Interactive strategies are sought that enhance the impact of the artwork; strategies that enrich the experience of the viewers, express better that which should be better expressed or allow for deeper, more thoughtful reflection. On this matter, little more can be said without falling into a discussion on the purpose of art, which is well beyond the knowledge of the author. Of one thing the author is sure: of all the qualities an artwork might potentially possess, none can be considered universally desirable, and this includes both immersion and interactivity. Any discussion, therefore, on the artistic benefits of immersion and interaction can only be made in the context of a specific artwork.

Using the flexible framework described above, a great number of different mappings and strategies were trialled in the search for an effective interaction paradigm for *La Dispersion Du Fils*. From this period of experimentation, much was learned about the nature of interaction within a shared immersive environment, including some results that run counter to preconceived notions. These results can be explained using the framework for presence presented at the beginning of this thesis, but before doing so, the model of interaction employed in Shanghai is very briefly described.

4.13.3 Shanghai Interface

The interaction strategy adopted for the Shanghai exhibition was, after much experimentation, decidedly simple. A 2 metre diameter circle was clearly demarcated at the centre of the AVIE arena, and this circle was used to trigger the transition to and from the Helix and the Circles. Whenever a viewer stood within this centre-spot, virtual passage along the Helix would steadily accelerate, and the viewpoint would return to the interior of the Helix, if it had not been already. If the viewer remained in the centre-spot for 30 seconds a maximum velocity would be reached, at which point the Helix would abruptly shatter into a million shards and disperse into a black void.[213] From this inky emptiness one of the 31 Circles would emerge. In this respect, the centre-spot functioned as little more than a button that could be held down by physically standing upon it. In addition, entering and exiting the centre-spot triggered rapid acts of mutation in the Helix.

[213] If they exited the centre-spot prior to reaching escape velocity, the virtual passage would rapidly decelerate.

In addition to this simple trigger, a number of the Circles possessed their own idiosyncratic interactive strategies. For example, in the Circle X.1 *L'Insulte Faite Au Paysage – Fioretti*, 18 films were arranged in a simple circle around the AVIE. As a viewer approach the screen, the closest film would react, zooming in from the distance towards the viewer and then floating just within the perimeter of the AVIE. If the viewer stepped backwards, the film would recede again into the distance, or if they moved sideways, another film would take its place. This extremely simple interactive strategy allows the viewers to select and watch any of the 18 films within the Circle.

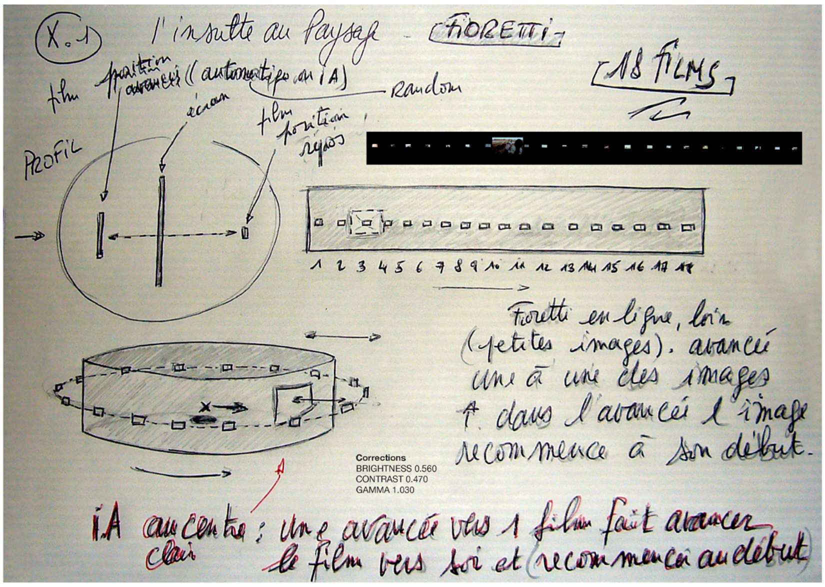


Figure 122 - Circle X.1 - L'Insulte Faite Au Paysage – Fioretti.

In other Circles, interactivity was employed less overtly. Here, the film-objects gently react to the presence of viewers with subtle alterations in trajectory or speed, showing attractive or repulsive motions or deviations in path.

4.13.4 Discussion

There is a prevalent and little challenged assumption that interactive experiences are inherently more immersive than their non-interactive counterpart.[214] The experiments conducted prior to the Shanghai exhibition, and the experience of the Shanghai exhibition itself, suggest this is not the case. Rather, with a multi-user

[214] This section compliments the discussion on interaction and immersion presented in Section 2.5.7.

immersive system such as the AVIE, and an abstract work of art such as *La Dispersion Du Fils*, it can be very challenging to find any interactive strategy that does not diminish presence for the viewers. There are a number of reasons for this.

The first is that, with a multi-user shared experience, it can be difficult to find strategies that give rise, for all viewers, to the perception of causality between their actions and perceived events. All of the strategies trialled that involved some aggregation or averaging of properties across multiple viewers failed to give rise to this direct perception of cause and effect, for although the viewer could sometimes perceive the effect of the audience as a whole on the virtual world, they would not perceive these effects as immediately and directly causally linked to their own actions. To repeat the position outlined previously: interaction contributes to presence only in so far as it facilitates active-perception of the virtual world, and for this there must exist a nomological and familiar relationship between efferent, reafferent and afferent cues. With a multi-user interface, these cues can appear to deviate from known patterns, as the actions of other viewers introduce what are perceived to be inconsistent and nondeterministic stimuli.

These multi-user strategies also vastly reduced the viewer's ability to *intentionally* influence the virtual world. Such 'democratic' interfaces are like any democratic system, in which the will of the individual is greatly subjugated to the will of the mass. Unless the audience enjoys a perfect unanimity of intention there will always be a certain level of frustration and disempowerment suffered by some, if not all, viewers. And if one belongs permanently to a minority, one may never perceive a relationship between one's intentions or actions and the unfolding events. These frustrations draw the conscious attention of the viewer away from the virtual world and onto themselves and other audience members, as they struggle to master a cumbersome interface.

A similar threat to active-perception and the perception of causality appears, even for a single viewer, when the mapping between the real world and the virtual world is too indirect, complex or anisomorphic. Although the actions of the viewer may numerically drive the virtual phenomena, if there is not a sufficiently evident and sufficiently immediate correspondence between the viewer's movements and perceived change, the two are unlikely to be directly perceived as cause and effect. That the mapping be isomorphic infers that cause and effect share certain qualities; not only must there be a perceived relationship between action and reaction, some aspect of the character or quality of viewer's actions must be reflected in the effected phenomena. David Rokeby, in his pioneering and exhaustive search for effective mappings between bodily motion and sound, encountered much the same problem:

> In the early days of "Very Nervous System" I tried to reflect the actions of the user in as many parameters of the system's behaviour as possible. I worked out ways to map velocity, gestural quality, acceleration, dynamics, and direction onto as many parameters of sound synthesis as I could. What I found was that people simply got lost. Every movement they made affected several aspects of the sound simultaneously, in different ways. Ironically, the system was interactive on so many levels that the interaction became indigestible. People's most common response was to decide that the sounds from the system were not interactive at all, but were being played back on a cassette deck. I found that as I reduced the number of dimensions of interaction, the user's sense of empowerment grew. Simplifying the language of interaction by reducing its variables let people recognize their impact on the system immediately. With repeated exposure, the user could handle and appreciate more nuanced levels of interaction. In time they could appreciate the flexible, expressive power I'd been trying to offer in the first place. (Rokeby, 1998)

Rokeby recognised that any sufficiently complex or novel interface requires a period of learning, if the user is to intend and then perceive correlation between their effects and actions. This presents a challenge to the publicly exhibited immersive artwork, such as *La Dispersion Du Fils*, in which not only are novel interface strategies often intentionally chosen, but a training period is often considered both practically and aesthetically undesirable. This results in the banishment of any interactive strategies that cannot be immediately grasped, understood and mastered by the average viewer. Rokeby understood the implications of this:

> As perpetual new users, we may be drawn inexorably toward simplistic systems, trading real power for an ever-evolving glimpse of some never-to-be-achieved potential. (Rokeby, 1998)

Whenever the connection between cause and effect is not obvious, the viewer's natural tendency is to actively search for causal connections, usually by exhaustively charting the territory of permissible actions and cataloguing their consequences. This explorative behaviour has an element of playfulness to it, as the viewer treats the interactive system like a puzzle that must be somehow solved. The emergence of this playful, explorative behaviour was observed repeatedly during the Shanghai exhibition. New viewers, upon learning that the AVIE could see them, would run about the arena in search of new ways to see their actions reflected on the screen. Only once the initial novelty had subsided, and the viewer felt confident that they had fully understood their causal powers and limits, were they more likely to settle into a more thoughtful relationship with the work.

This period of familiarisation presents a particular problem for a shared, multi-user experience with no beginning or end, such as *La Dispersion Du Fils*. The exploratory play of the new user can be very disruptive to the other members of audience who have already moved on to more reflective states of participation, and as viewers can enter or exit the AVIE at any time, there will almost always be someone within the AVIE for whom the experience is novel. This touches on a broader challenge inherent in group interfaces: they demand of the viewers a certain uniformity of personal experience and personal evolution.

Certainly these problems can be overcome. Artworks can incorporate periods of induction during which the audience learn the interface and the behaviour of one another. This is the approach taken, for example, by CREW in their work *Terra Nova*, where instructional elements are written into the performance (see Section 2.6.2). This approach was not, however, considered compatible with the never-ending character of *La Dispersion Du Fils*. In any case, unfamiliarity of the interface was not the only impediment to effective interaction.

Inviting the viewer to consciously and intentionally interact with the virtual world demands of the viewer some understanding of their role in the unfolding narrative. In many interactive experiences, this role-playing is one of the principle sources of enjoyment; temporarily becoming a race-car driver or interstellar pilot is a significant part of the allure of computer games. With *La Dispersion Du Fils*, however, no such role is offered the audience, and nor are they invited to imagine one for themselves. Rather, they are invited into a state of entrancement, a state of pure physical perception. As a consequence, the actions and intentions of the audience within the universe of *La Dispersion Du Fils* are entirely meaningless.

Demanding an audience exercise their powers of causality while withholding from them any explicit roles or goals is observed to have a single outcome – the viewers make a *game* of it. Upon learning of their causal powers, the audience feel immediately compelled to act, and if the virtual world offers no clues as to the 'correct' behaviour, the viewers will simply act incoherently rather than not act at all. The emergence of game-playing behaviour - behaviour which has very little to do with the virtual world depicted in *La Dispersion Du Fils* - was observed during the Shanghai exhibition.

The most compelling explanation for the diminishment of presence with interaction, however, can be traced to perceptual conflicts introduced or aggravated by acts of interaction. By demanding physical movement of the viewer, the interface exasperates the flaws inherent in the reproduction of the light-field: lack of motion parallax, the poor resolution, the boundary between the real world and the virtual. As discussed previously, the vision of the light field offered by the AVIE is distorted, and the nature of this distortion varies with the position of the viewer within the AVIE. It has been observed that for the stationary viewer, regardless of their viewpoint within the AVIE, these distortions are largely imperceptible. It is when the viewer is in motion that they become evident, and disruptive. By inviting the viewers to move about in the hope of asserting their influence on the virtual world, these flaws in mediation are brought to the fore.

4.13.5 Abandoning Interaction

There are many possible reasons for introducing interaction into an artwork. In the case of *La Dispersion Du Fils*, interaction was introduced to heighten immersion. In practice, however, interaction, in the form described above, was found to significantly diminish immersion. Further, it was clear that such reductions in immersion dramatically reduced the power of the artwork as a whole. Interaction seemed to offer no benefits that might justify or counterbalance such a loss, so following these initial experiments and the Shanghai exhibition, it was decided that all subsequent versions of *La Dispersion Du Fils* would be non-interactive.

It is worth repeating that these results are naturally highly contextual, and can only be judged in the context of *La Dispersion Du Fils*. The interaction strategies described above, for example, have been known to work perfectly well in other situations.[215] However, it can be safely said that while interaction may have the potential to enhance presence, it can very easily destroy it, for any of the reasons presented above. Further, it would appear highly beneficial, to the artist or designer, that the role of interactivity in the overall experience always be made explicit, whether it be an intended increase in engagement, better information recall, heightened immersion, or greater physical activity (for example), for only then may the effectiveness of one interactive strategy be judged meaningfully, and the use of interaction for the sake of interaction be avoided.

[215] With *Rollerball*, an AVIE project created under the supervision of the author during the iCinema Studio 2009 seminar, an entire game was constructed around the interactive strategy of using the centre of the group to roll a giant ball through a virtual maze.

4.14 LANDSCAPES

Following the Shanghai exhibition, development of a new series of Circles began, during which a different approach to extra-Helicular phenomena emerged. The Circles had largely been designed as polyptychs; virtual exhibition spaces where one could regard the films in their entirety. And in this respect, they fulfilled their function. They did not, however, tend to provoke in the viewer a sense of passage. Within the Circles, the action and events were entirely viewer-centric, with the movement and orientation of the films firmly placing the viewer at the centre of the scenario. That is, the Circle film-objects always appeared to travel towards the viewer, rather than vice versa, and even though, numerically speaking, the viewer's progress continued unabated, neither the Helix nor any other structure was at hand to reveal this motion.[216]

Further, the Circles were entirely composed of free-floating structures, without ground, scaffold or firmament, and while they often exhibited a coarse collective motion, it was always visibly an emergent one, as if it were a learned behaviour that ran against their true nature. In giving the film-objects a certain organic animacy, rather than moving them in rigid formation, the viewer was deprived of any structure that might be perceived as a perceptual rest frame. This was further compounded by the black void perpetually visible behind the free-floating film-objects, which clearly established the film-objects as foreground. In short, the Circles offered neither a sense of ego-motion, nor a perceptual rest frame, and the result of this was that the viewers, and the film-objects, were temporally returned to the real-world space of the AVIE. This is not necessarily to say that presence had been diminished, but that a switch in reference frame had occurred, and now rather than perceiving themselves as moving through a stationary virtual world, the viewers perceived virtual objects moving through the stationary real world, a perceptive switch that tended to disrupt the sense of endless voyage.

[216] Recall that the Helix is never visible at the same time as the Circles: the Helix shatters and disperses to reveal a void, and from the depths of this blackness the Circle elements approach.

And so, in the light of these observations, a somewhat different approach to these events and phenomena was adopted. First, the absence of the Helix was addressed. From now the Helix would always be present, and these extra-Helicular happenings would be presented as something that travelled to or with the Helix, or as places to which the Helix delivered the viewer, or as internal happenings, within or as part of the fabric of the Helix itself. Most importantly, the Helix would never be absent, so that it might at all times provide a reference frame, a sense of motion and a sense of continuity. It would always be at hand to lend a measure of scale, facilitating the display of monumental architectures and structures that might completely engulf the Helix, or render it insignificant against a vast backdrop or amidst a vast terrain. As this new series of events and phenomena - endless plains, oceans or terrains, or infinite fields of images, or structures that entirely envelop the Helix – can no longer be described as Circles, they shall be referred to as extra-Helicular phenomena, or E.H.P.

This movement towards endless or gigantic structures was partially a response to a very immediate and basic challenge posed by the panoramic theatre: a 360° vista presents a lot of empty space to fill. Any object presented in the foreground yearns for a background, and the provision of visual backgrounds is vital to the perception of rest-frames and ego-motion.

In a typical exhibition of the work, E.H.P are programmed to appear with a slow, intermittent rhythm, separated by no more or less than 5 and 10 minutes in time, and arriving in a more or less random sequence. Each phenomenon would unveil itself in one of two ways. For phenomena with finite size, the Helix would temporarily and abruptly suspend its wayward meanderings to unfurl and stretch out with sudden purpose and direction. As the Helix, and we with it, gather more and more speed, the object of desire reveals itself, at first only faintly visible in the distance, but growing ever larger as the Helix grows and decays towards it, with ever-increasing urgency. Before long, the Helix has delivered the viewer to the heart of the structure, and only then it resumes its habitual, labyrinthine motion.

For phenomena of unbounded dimensions – the landscapes – the viewers emerge from a long spell within the Helix to find the new world already surrounding them, not unlike passengers surfacing onto the deck of a ship after making port during the night. Here, it is not uncommon for the sound of the landscape, audible through the skin of the Helix, to forewarn the passenger of their arrival in some foreign new land.

Presented below (in no particular order) is a brief technical and thematic account of the half dozen E.H.P that reached sufficient maturity for exhibition. It should however be noted that the thematic interpretations offered here are entirely the authors. They are interpretations made in hindsight, and were never part of any design or discussion shared between the artists, and are therefore not necessarily reflective of the views of the Bruyère, Varas or Arredondo. They are offered not so much to help convey to the reader the experience of witnessing these structures within the AVIE, but to draw attention to an important aspect of the work: the capacity for immersive virtual worlds or structures to convey complex themes without recourse to symbols. And neither the very precise engineering behind the visions nor the immediate, raw and visceral mode of presentation, prevent an immersive work like *La Dispersion Du Fils* from possessing a certain poetry. The work invites interpretation and the assignation of meaning no less than literature or cinema, and any fears that immersive art must somehow give way to a less poetic, less reflective or less 'critical' experience are without cause; typical reactions to an emerging medium and nothing more.

4.14.1 $CaMg[CO_3]_2$

> The pack, eager for its prey, swept over the rocks and crags, over unapproachable cliffs, through places where the going was difficult, and where there was no way at all.
>
> *Metamorphoses.III.225-229 (Ovid & Innes, 1955)*

In 2007 Bruyère and the LFKs were invited by the ZKM Institute of Visual Media to envisage a new application of their panoramic video camera. The panoramic camera, one of many developed at the ZKM, was composed of a ring of 15 digital video cameras, their fields-of-view slightly over-lapping to allow the subsequent stitching and blending of the individual video streams into a single seamless 360° panoramic video.

The result of this commission was *$CaMg[CO_3]_2$*, and the central element of the work was a spectacular, soaring flight through the valleys and along the ridges of the Dolomite mountains, captured by mounting the panoramic camera-array on the underside of a helicopter.[217] The footage from this project would be re-purposed for *La Dispersion Du Fils*.[218] However, rather than simply re-project it as a panorama, as it had been in *$CaMg[CO_3]_2$*, the film would be integrated into the virtual world of *La Dispersion Du Fils*.

Figure 123 - $CaMg[CO_3]_2$ *still frames (LFKs, 2007).*

Initial experiments followed the concept of 'film as surface,' in which the cylindrical film was mapped to the surface of a geometric body. A number of different geometries and topologies were examined, and those sharing the topology of a cylinder proved natural and obvious candidates. Distorted, bent or folded, but maintaining the essential topology of the cylinder, cylindroids were inflated to enormous dimensions, tens of kilometres in size, such that they might engulf the Helix in its entirety. The result is a form of hollow planet, bounded by an event horizon, but with an aperture through which light, or a traveller, might arrive or leave. The sporadic appearance of these inverse-planets forms part of the greater cyclic narrative of *La Dispersion Du Fils*. The Helix, with a sudden expression of purpose, lurches towards the faint bluish orb barely visible in the distance, and as the Helix rushes forward with ever increasing fervour, the

[217] Dolomite = calcium magnesium carbonate = $CaMg[CO_3]_2$

[218] With the kind permission of Bernd Lintermann, director of the ZKM Institute for Visual Media and co-creator of the original *$CaMg[CO_3]_2$* panorama.

viewer becomes aware of first the distance and then the size of the looming planetoid. Only once the Helix has reached its destination - the optical centre of the 'eye of the world' - does it relent from the chase, slowing to a gentle, silent flight above the wooded slopes and barren cliffs of the pictured mountain ranges.

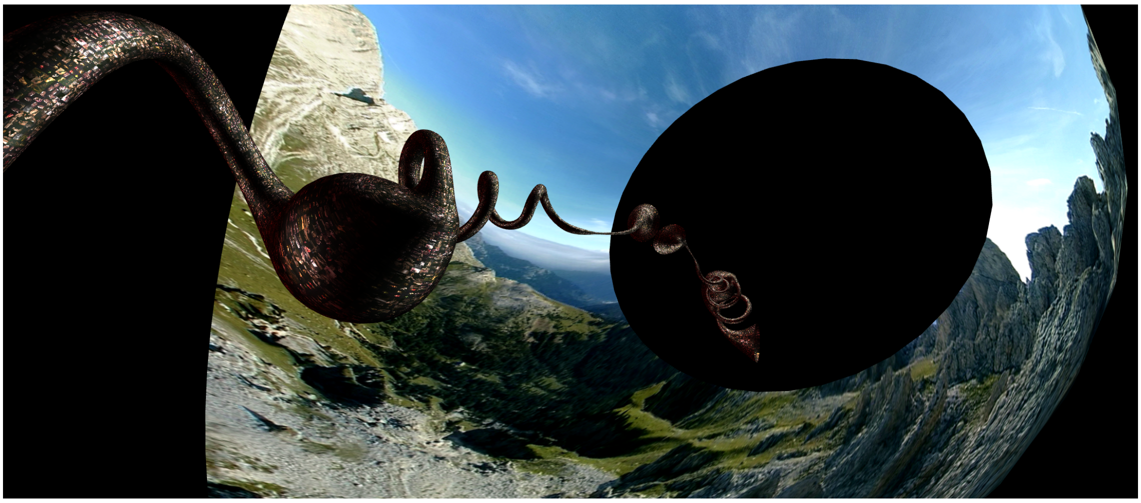


Figure 124 - The Helix and the 'eye of the world.' Screenshot.

Such is the habitual return of the hounds to Gargaphie, the ancient valley where Actaeon first held Diana's gaze. It is a return to the mountains where they once hunted with their master, mountains that by knowing them so well they were able to hunt him down and end his life. They return in search of their master, in search of a sign, or a resurrection. They return in pursuit of their former blindness, hoping to reclaim their wordless nescience and win back their thraldom. They return, above all, because they are creatures of habit; in a ritual observed for long forgotten reasons.

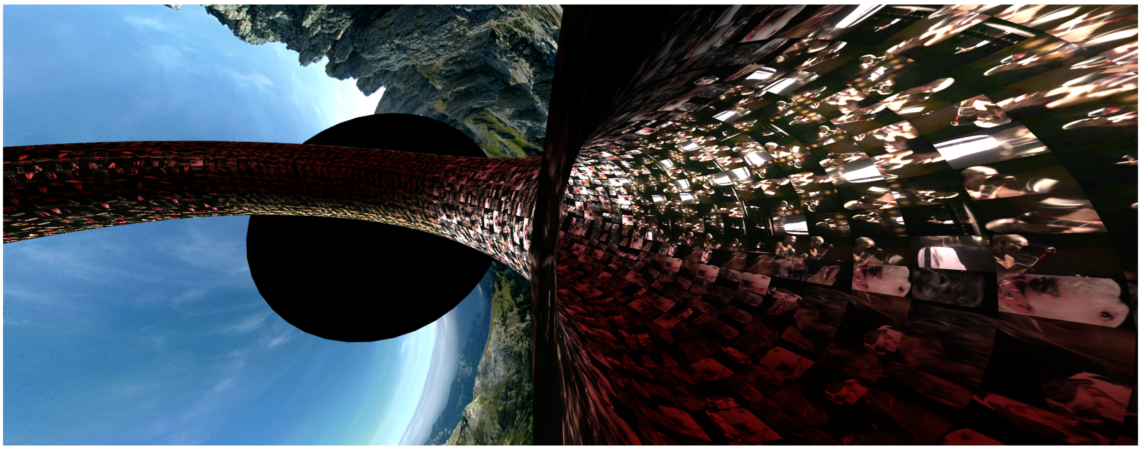


Figure 125 - Helix within the panoramic video. Screenshot.

It is, in fact, an imaginary journey, occurring nowhere but within the minds-eye, a practiced remembrance, betrayed by the superficial, skin-like nature of the rolling mountains; they are, after all, *images*. These mountain groves no longer exist, save in memories of memories, and even if they did, they have long been lost to these vagabond exiles. It is, then, a periodic withdrawal from the world, a temporary refuge found in an imaginary inner-world, where all was as before, but may never be again.

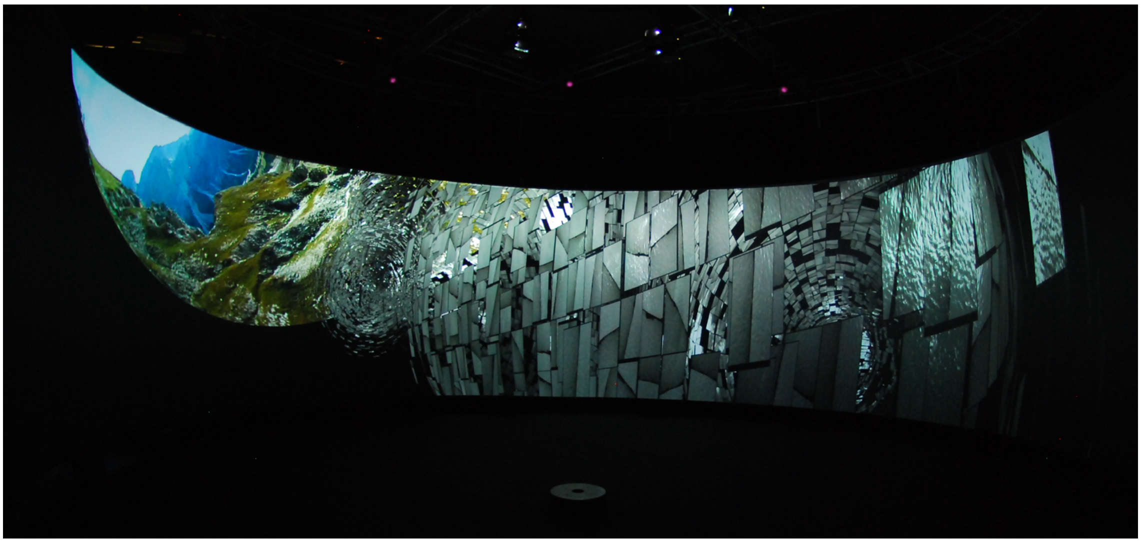


Figure 126 - The Helix approaches the panorama. Photo in AVIE.

Implementation

The panoramic video has resolution of 3480 x 480 pixels and runs 29 minutes 40 seconds, and in its compressed (mpeg2) form, occupies 6.2 GB of disk space.[219] Smooth, jitter-free playback of a video of this resolution was the significant challenge, for at the time *La Dispersion Du Fils* was in development, the transfer, decompression and display of this data while simultaneously computing and rendering the Helix, was not a trivial task. As the video was mapped to the inner-surface of a sphere with no fixed orientation relative to the AVIE, the video could not simply be segmented into 6 sectors and divided among the render-nodes. Each render-node instead required instant access to any or all parts of the panorama. As such, on all machines the entire movie-frame was streamed from disk, and decoded into YUV images in RAM. Then, only the visible

[219] Uncompressed, in raw RGB form, the panorama occupies 230GB, equivalent to 132 MB/s.

portions of the panorama were transferred to VRAM, where they were transformed into RGB images, their 'MIP-maps' calculated, and made ready for display.

A great deal of work was required to ensure that the streaming from disk and decompression did not disrupt the display of the Helix, by using multiple threads to read ahead and decompress. This presented delicate inter-computer timing challenges, for while the frame-rate of the panoramic video was decoupled from the render frame-rate, the render-nodes must still present the same frame of the video at exactly the same time.

4.14.2 Forest

There is a tale that when Jupiter saw Diana walking unclothed, he ordered Mercury to make her a garment. But try as he might, he could not make one that fitted, for the moon Goddess was continually changing her size.[220] The story reflects Diana's role as the figure of increase and decrease, of growth and decay, of shifting tides and changing mood, of the perpetual change that underlies the natural world. This facet of the deity is what Nuccio Ordine refers to as the "infinite mutability" of Diana; the paradoxical "constant in everything" (Ordine, 1996, p. 88).

This infinite mutability is nowhere better illustrated than in Diana's natural dominion: the forest. Here, the infinite multiplicity of forms, the unending variation of self-similar structures, from giant trees to the smallest fronds, is made all the more splendid for its refusal to repeat - its infinite mutability. A forest, were it to extend from horizon to horizon, would contain countless forms without recourse to a single repetition. And with it an unstayable process of growth and decay ensuring that all is in motion, even when still.

In this landscape, the viewer accompanies the Helix through an infinite field of cubic forms in a constant state of flux. Images of a journey through a forest is mapped to the surface of the cubes (the 76 films of *Paysage - Foret*), and the arrangement and motion of cubes is defined by a large number of continuous parameters. The cubes are,

---

[220] The story is depicted in the *Hieroglyphicorum* of *Caelius Augustinus Curio,* and recounted in Valeriano (1556).

therefore, designed expressly to evolve as their parameters trace a meandering trajectory through the space of all states.

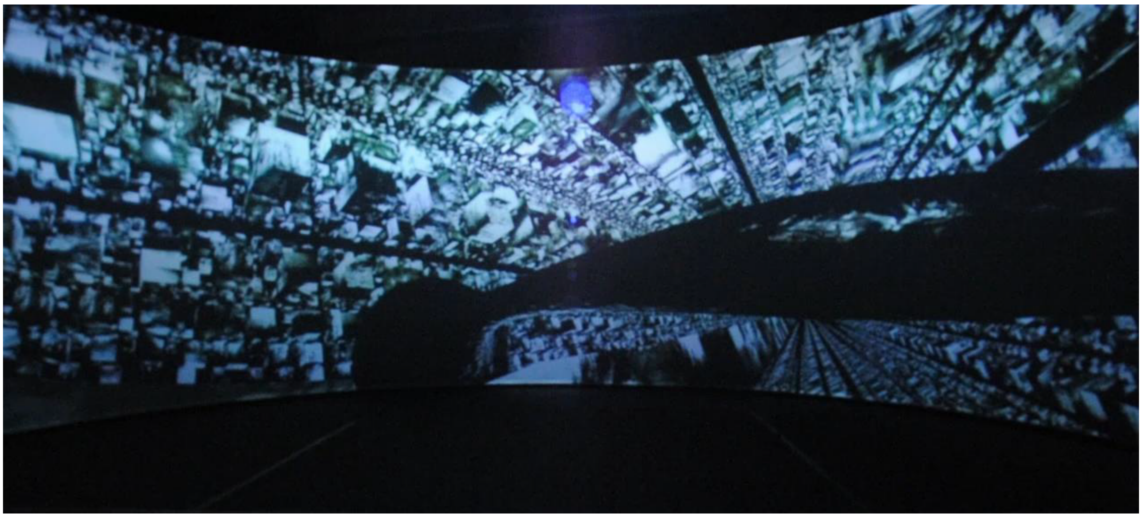


Figure 127 - Forest of Cubes. Photo of early prototype in AVIE.

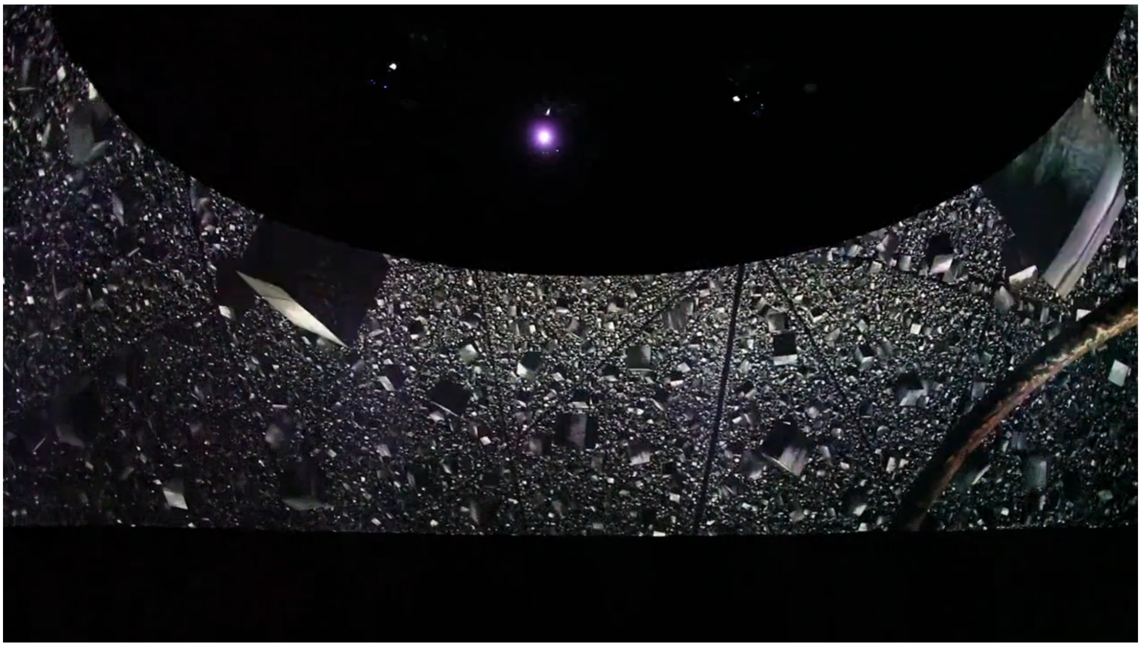


Figure 128 - Forest of Cubes. Still image from video captured in AVIE.

4.14.3 Larynx

> clamare libebat:
> 'Actaeon ego sum: dominum cognoscite vestrum!'
> Verba animo desunt; resonat latratibus aether.
>
> Ovid, *Metamorphoses III.229-231*

Verba animo desunt. Words fail the will. Actaeon is surrounded by the pack and all hope of escape is lost. He longs to cry "I am Actaeon. Recognise your master!" but his thoughts are denied words, and the air resounds with barking.

Of his transformation into a deer, it is this alone - *the loss of language* - that deprives him of his identity and humanity, and ultimately, his hide. It is, above all, his inability to speak the names of his killers, or to name himself, that brings upon him such savage and unrepentant slaughter; a murder in every sense, but for his wordlessness. (Is this truly all that distinguishes the murder of a man from the slaughter of an animal – the spoken word?) A murder made all the more horrific for the unaltered state of mind within, for "*mens tantum pristina mansit*", "only his mind remains unchanged."[221] This is where the nightmarish quality of Ovid's Actaeon resides: in the horror of "the mind unmoored,"[222] a mind without the means with which to reveal itself.

The spoken word lies at the centre of this third external event. The viewers, temporally ejected from the belly of the Helix, find themselves silently gliding across the surface of a gently undulating, glistening surface. Looking about, it becomes apparent that they are bound to the inner-surface of some vast chamber, the opposite sides of which a barely visible several kilometres across the AVIE. In the centre hangs their familiar companion the Helix, silently curling upon itself. They have been cast adrift, but not completely abandoned.

---

221 Ovid, Metamorphoses III.203

222 Cavell (1979, p. 472) quoted in Bruns (2011, p. 32)

Figure 129 - The Helix floats within a vast undulating chamber.

Suddenly the skin of the chamber tenses and momentarily draws taught, before a sound, deep and reverberant, fills the enormous chamber, sending the chamber's membrane into paroxysms and contractions. Silence and stillness follow, and then a second sound erupts from the walls of the chamber. It is clearly a voice, but an alien and unintelligible one. The walls of the chamber undulate and convulse to words with undecipherable meaning. The Helix, still twisting silently in the centre of the chamber, also moves with each word, visibly pulsing and twitching with each vowel and consonant. Across the surface of this resonating chamber, images and the written word drift and strain, without ever quite revealing their meaning.

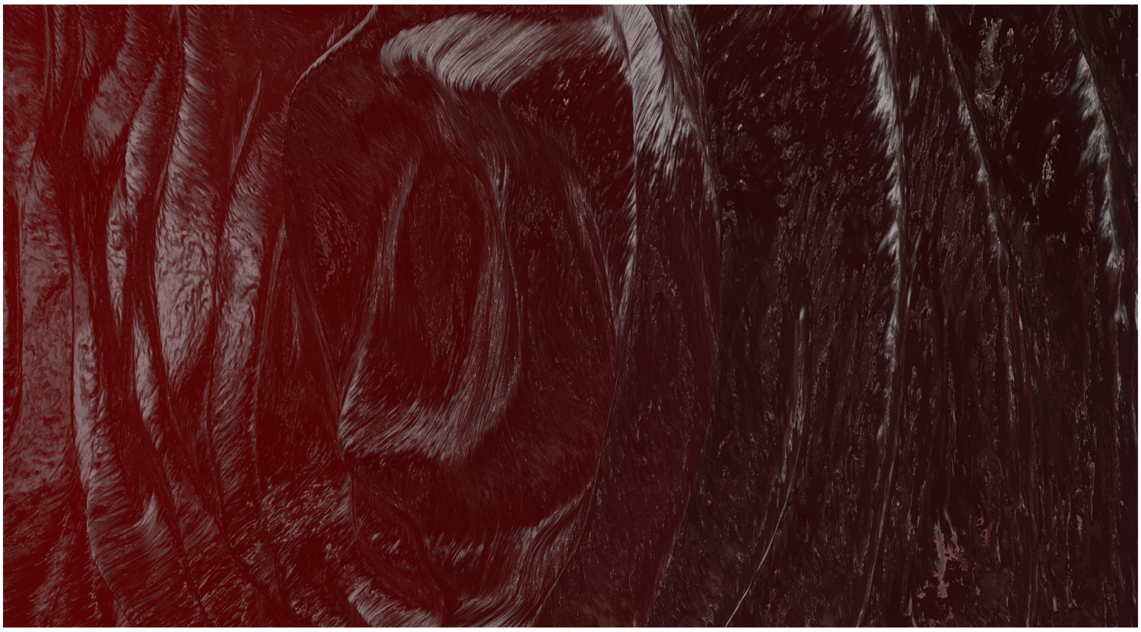


Figure 130 - The walls of the chamber undulate with the soundtrack.

Soon, it becomes clear that the voice is not speaking, but singing. And while the meaning of each word might be lost, the sentiment is not, for it is clearly some terrible lament, a dirge in a long forgotten language. But who is singing? In the *Cymbalum Mundi*, Bonaventure des Périers tells of Pamphagus and Hylactor, the two dogs who devoured Actaeon's tongue, claiming for themselves the faculty of speech. (One man's loss is another man's gain.) Des Périers describes their chance meeting by the roadside, where, having revealed their gift to no one, they agree that it is better to stay silent, to return to the life of a dog and forget all that might be said with the human tongue. Is this, then, the lament of a dog we are hearing?

Figure 131 - Pamphagus and Hylactor illustrated in des Perier's Cymbalum Mundi, 1537. "Dialogue between two dogs, who having eaten the tongue of their master Acteon, giving them the faculty of speech, read the letters that they find on their path."

Implementation

The chamber is formed by a sphere, (subdivided invisibly into segments to permit efficient culling), whose radius is distorted with a 4D dimensional multi-octave noise field. The frequency and amplitude of each octave is tied to the output of a spectral decomposition of the soundtrack, extracted in real-time with a fast-Fourier transform.

The accompanying video, *Elements - Paysage II,* is mapped to the diffuse albedo of the inner-surface of the sphere, and spatially distorted through another noise-field, which too is coupled with the spectrum of the song. The skin of the chamber is bumped with the height-map derived from the antlers of a deer, and the strength of this is modulated according to the degree of undulation, to crudely simulate the stretching of a membrane. The viewer travellers gently over the inner-surface of the sphere, while the Helix orbits focal point in the centre of the chamber. The Helix is programmed to react to the sound also, by coupling the strength of its shape deformers to the output of the frequency analysis. The soundtrack to the video contains a sole male voice, singing in an unrecognisable language. The video itself contains subtitles, which are visible across the surface of the chamber, but always illegible. For the record, they read:

and death returns
without anyone knowing
what has been seen.
death returned to language… through its holes
is it serious? is it a crime??
no, absolutely not, it's good.
also, as always,
one should not be born
because only this wins over all words.
but if one is shown the light
less damage is done by returning
to where one came from
and the sooner the better…
for everyone.
sing, Josef, sing…

4.14.4 Sirius the Dog Star

Seated in the jaws of the constellation *Canis Major* lies the Dog star, largest and brightest of all stars in the heavens. The Greeks called it Sirius, the "scorcher," for its heliacal rising marks the coming of the Canicula, the dog-days of summer when the Sun burns with doubled purpose. "No star comes on mankind more violently or causes more trouble when it departs," wrote Manilius (1977, p. 34), "it barks forth flame, raves with its fire, and doubles the burning heat of the Sun. When it put its torch to the earth and discharges its rays, the earth foresees its conflagration and tastes its ultimate fate."[223]

"All living things seek alien climes and the world looks for another world to repair to; beset by temperatures too great to bear, nature is afflicted with a sickness of its own making, alive, but on a funeral-pyre: such is the heat diffused among the constellations, and everything is brought to a halt by a single star. When the Dog star rises over the rim of the sea, which at its birth not even the flood of Ocean can quench, it will fashion unbridled spirits and impetuous hearts; it will bestow on its sons billows of anger, and draw upon them the hatred and fear of the whole populace." (Manilius, 1977, p. 316).

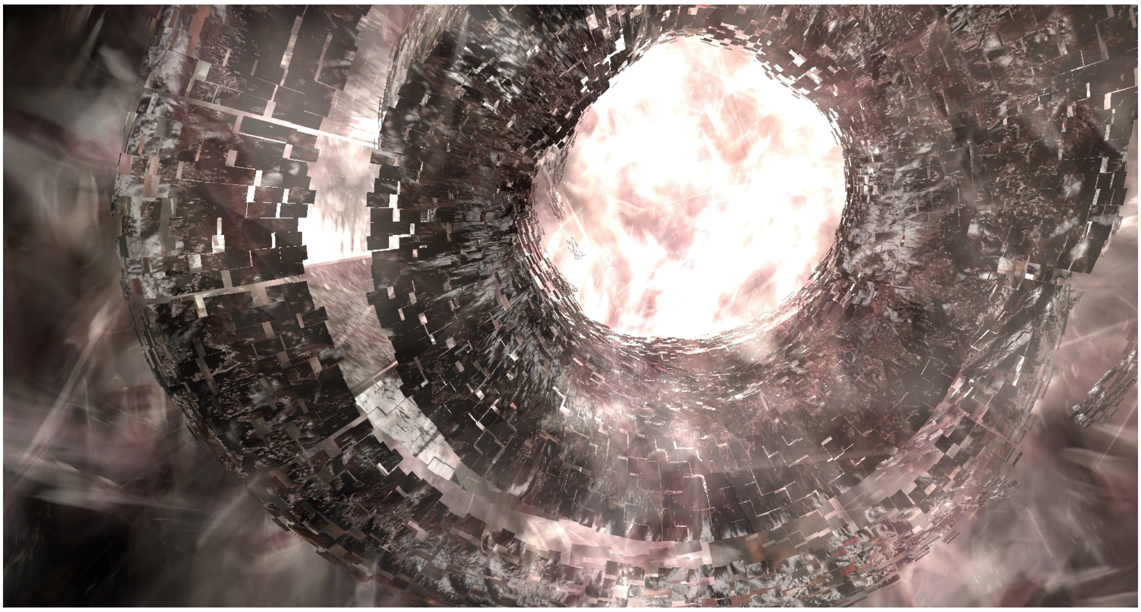


Figure 132 - Canicula I.

"Some authors say, from Hippocrates and Pliny, that the day this star first rises in the morning, the sea boils, wine turns sour, dogs begin to grow mad, the bile increases and irritates, and all animals grow languid; also that the diseases it usually occasions in men, are burning fevers, dysenteries, and phrensies. Hence the Romans sacrificed a brown dog every year to Canicula at his first rising, to appease its rage." (1810)

[223] A reference to *ekpyrosis* - the periodic destruction of the cosmos in which all things are returned to the primeval fire, pending renewal.

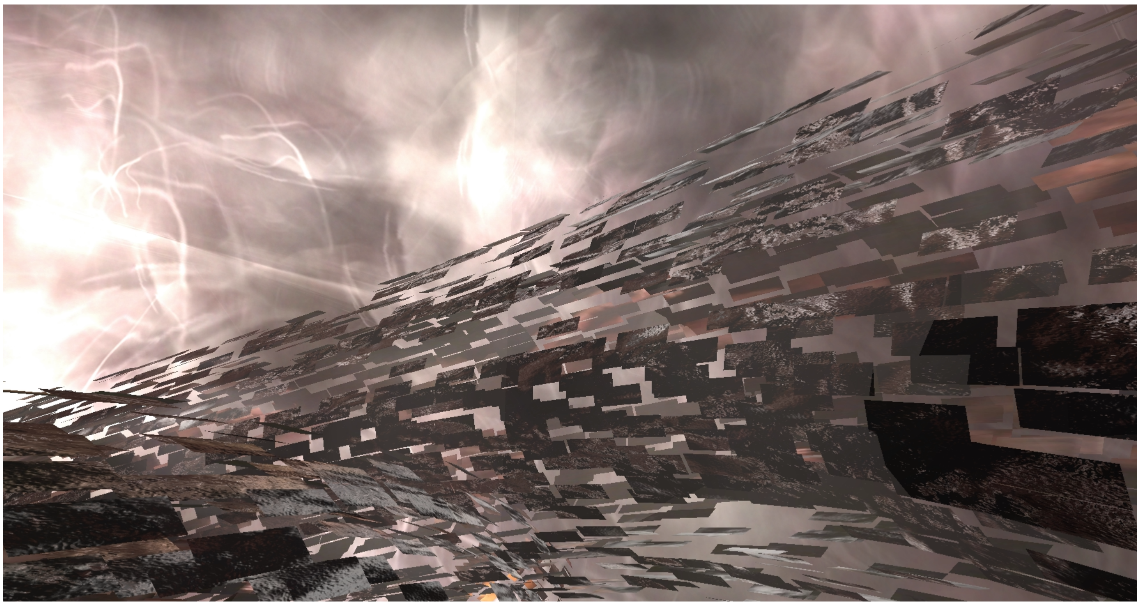


Figure 133 - Canicula II.

4.14.5 Elements

In some accounts of Actaeon's demise, Diana, after turning him into a stag, turns to Lyssa the spirit of madness, rage and rabies, and implores her to set the dogs upon on their beloved master. This is a misunderstanding. Lyssa, it is true, was present, but it is not the dogs she touches, but Actaeon himself.

> I stretch my thoughts to the sublime prey, and these springing back upon me, bring me death by their hard and cruel gnawing.
>
> *Bruno, Fourth Dialogue, Gli Eroici Furori, Bruno and Memmo (1964)*

The transformation of the hunter into the hunted, condemned to an eternal autophagia of self-pursuit. What is madness, if not this?


Figure 134 - Elements I.

In *Elements*, an alien landscape is depicted. A voice resonates across the plain, barking indecipherable instructions. The terrain shifts, pitches and pulses with a form of sentience.

> The dogs that know not their own master are mad dogs; […]
> There is […] the stag-man, attacked by his own bitches,
> with Lady Madness driving them on.
> Sweet heavens, do not make me mad.
>
> *(N. O. Brown, 1972)*

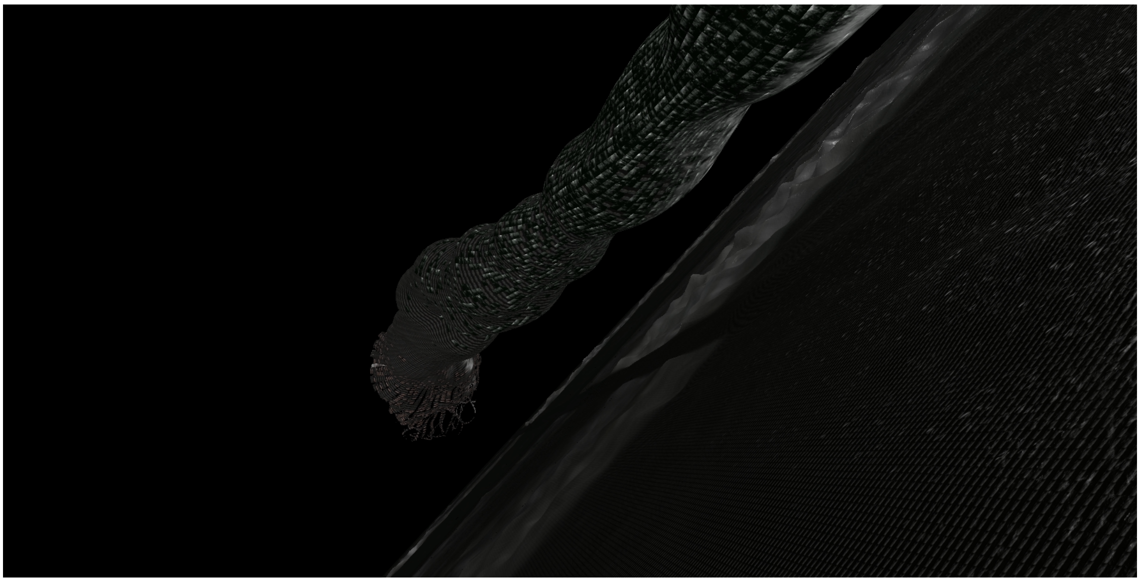


Figure 135 - Elements II.

4.14.6 Anabasis

mons erat infectus variarum caede ferarum

Ovid, *Metamorphoses.III.143*

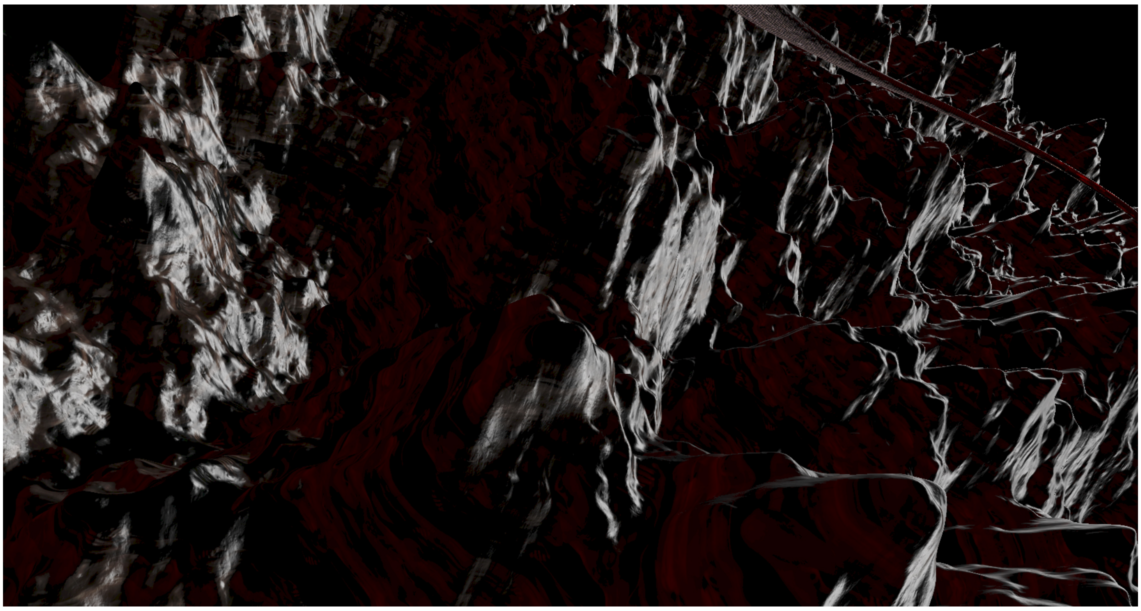


Figure 136 - Anabasis I.

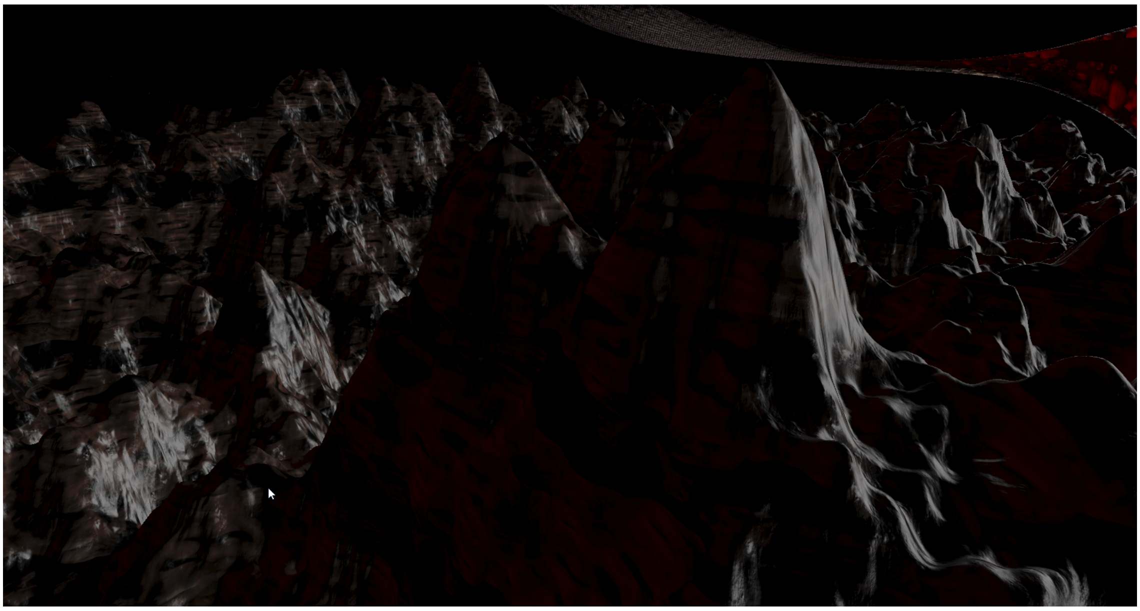


Figure 137 - Anabasis II.

Gray cliffs,
    and beneath them
A sea
Harsher than granite,
    unstill, never ceasing;

Ezra Pound, *The Coming Of War: Actaeon*

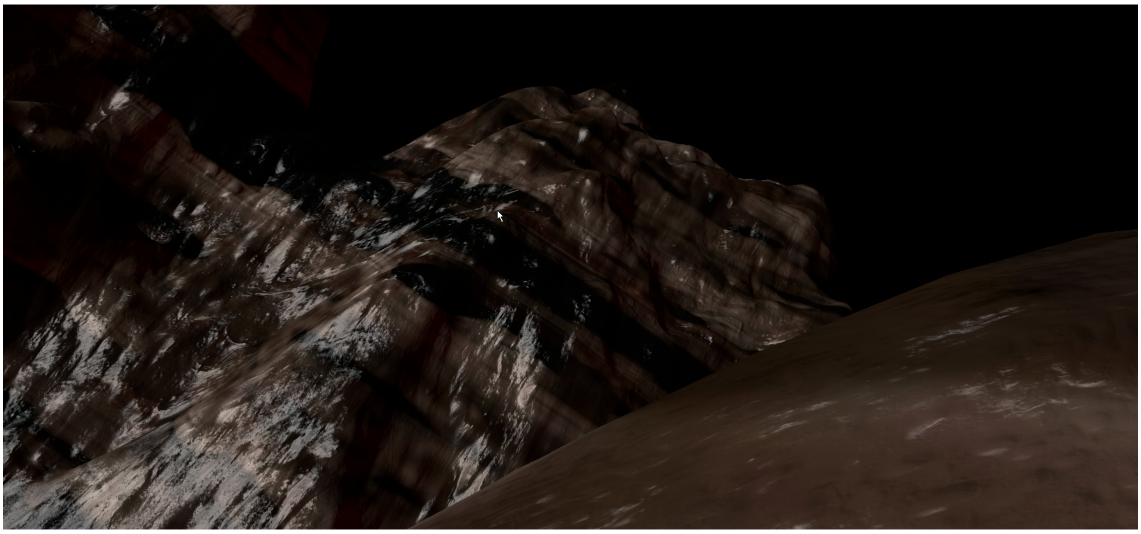


Figure 138 - Anabasis III.

4.15 MATERIALISATIONS: FROM THE VIRTUAL TO THE REAL

> Sculpture is this: discovery, by moving the body, of an ensemble of signs deployed in space; signs inescapably transformed by our encircling gaze.
>
> Jean Michel Bruyère.[224]

It is difficult not to draw parallels between this definition of sculpture and the active-perception theory of presence. To regard a sculpture is to physically act, and it is only to the moving observer that a sculpture reveals itself. By moving in space, different aspects are revealed, alignments and correlations created, and connections formed. This is the quintessential difference between an image and a sculpture.

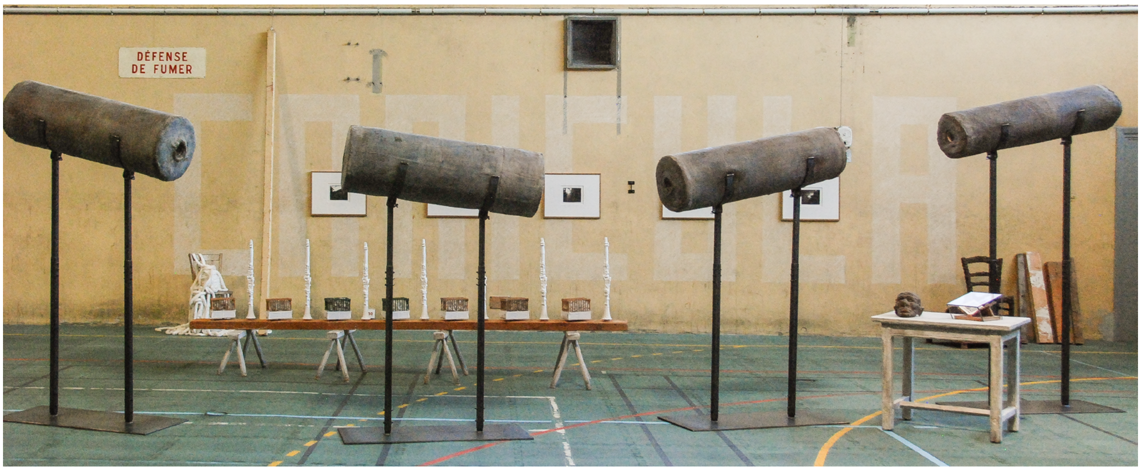
In the version of *La Dispersion Du Fils* presented at the 2011 Avignon Theatre Festival, France, the fictional world breaches the perimeter of the AVIE and erupts into the material world. Here, the AVIE was embedded within a constellation of sculptures and installations, projections and images, all created by LFKs. Installed within an old gymnasium whose floor - overlaid with the lines and markings of various sports, each competing for territorial dominion - is evocative of the geometric forces and trajectories dictating the virtual world.

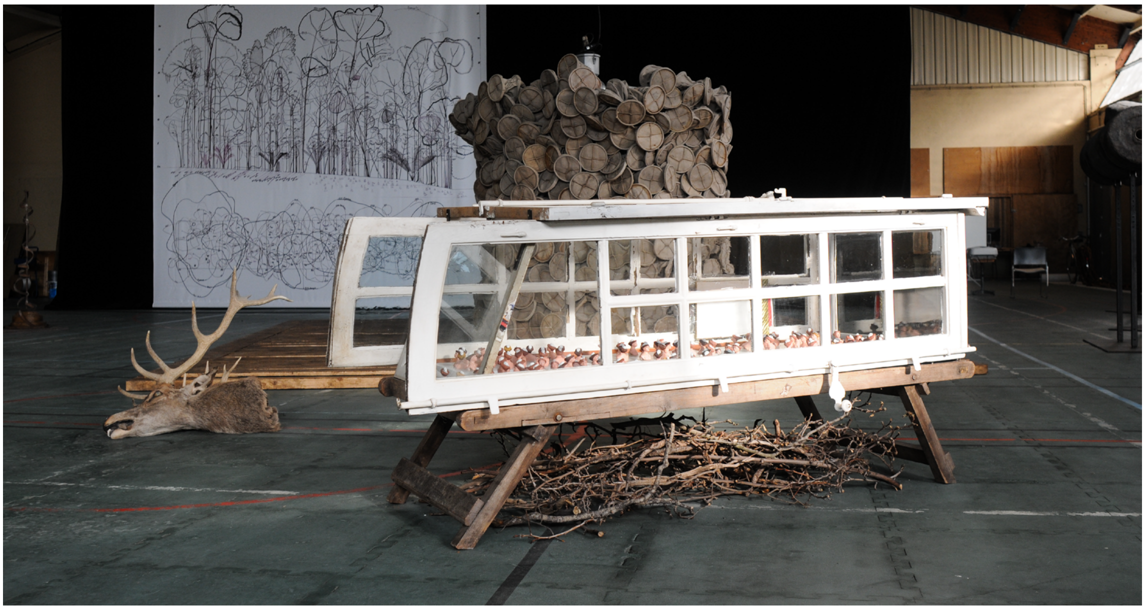
Twenty five distinct works surround the AVIE, including the severed head of a hart, an eviscerated television, an early print of Bonaventure des Périers' *Cymbalum Mundi* and a bicycle with a 50ft intestinal wedding train. The words LYSSA + HYALE = CANICULA writ along one length of the hall.[225] A small cinema shows the LFKs films *L'Insulte Faite au Paysage* (2006) and *Prima Stanza* (2005) in their entirety, while another installation features the voice of Francis Hallé, the French botanist celebrated for sailing the vast rainforest canopies of Amazonia in an inflatable sky-raft, propounding the superiority of the kingdom of plants over that of the animal.

---

[224] *«La sculpture c'est cela : la découverte par le déplacement du corps d'un ensemble de signes déployés dans l'espace et se transformant toujours tandis que le regard l'encercle.»*

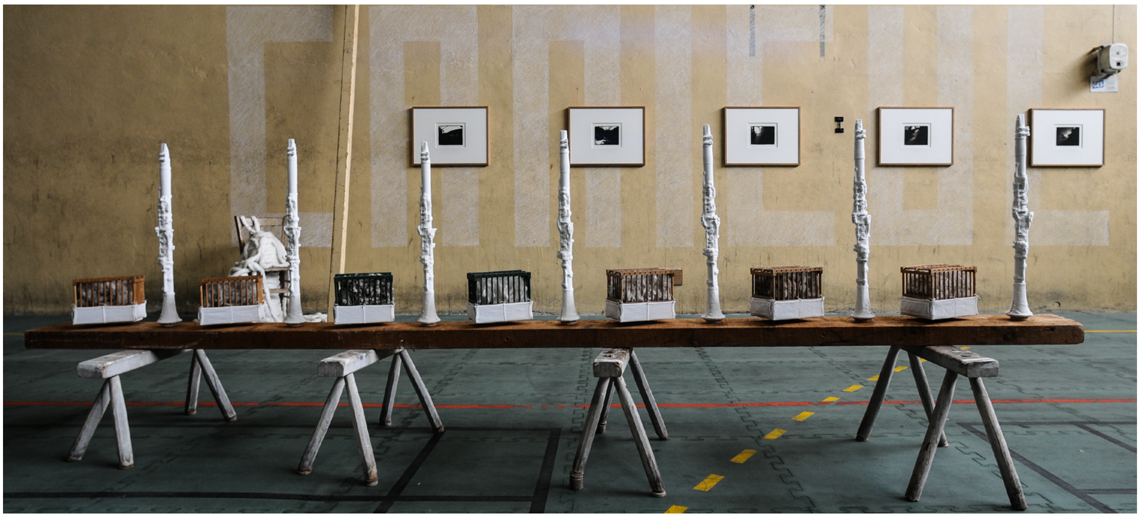
[225] Lyssa was the goddess/spirit of mad rage, frenzy and rabies. In some accounts of the death of Actaeon, it is Lyssa who sends the dogs into their blind frenzy. Hyale was one of the Oceanid Nymphs name in the retinue of the moon goddess Diana.


DÉFENSE
DE FUMER




DÉFENSE
DE FUMER


CYMBALUM
MUNDI.
DIALOGUE IV.
LES CHIENS
D'ACTEON.
ARGUMENT.




DÉFENSE
DE FUMER


in dog
we trust

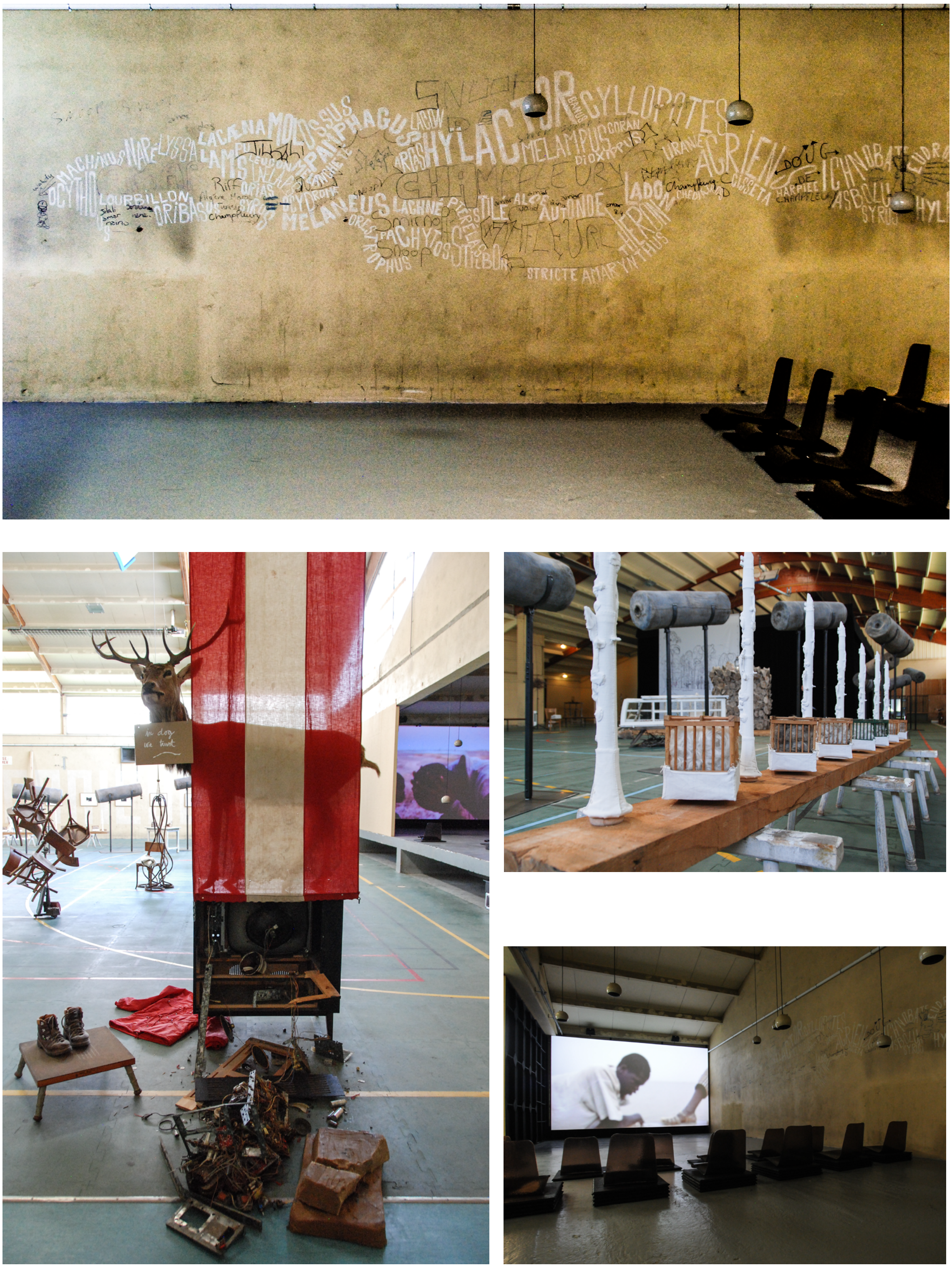


Figure 139 - La Dispersion Du Fils - 20 photos from the Festival d'Avignon 2011.

5. CONCLUSION

This thesis began by asking of the relationship between the panorama and immersion and presence. This question has been answered thrice, in three bodies of work.

The work began with a theoretical account of immersion and presence, based on an understanding of how we see the world. Drawing on J. J. Gibson's ecological optics, a picture of presence as the act of creating and detecting invariant structures in multi-modal stimuli. This notion of presence was arrived at by studying a host of perceptive phenomena including active-perception of the plenoptic function, degrees of freedom of vision, multi-sensory binding and cross-modal interactions, ego-motion, vection and perceptual rest-frames, interaction and the perception of causality. Special attention was given to the perception of images and the destructive effects of the image on immersion.

The second body of work concerned the AVIE; a re-incarnation of the panorama constructed with contemporary VR technologies. By employing omnistereo, spatial sound and interactivity, an effective, robust and highly versatile solution to the problem of immersing multiple viewers in a single shared space has been demonstrated.

Finally, the artwork *La Dispersion Du Fils* was created specifically to explore the immersive and aesthetic potential of the AVIE system and the panoramic form in general. The work explores the immersive potential of all of the elements of presence touched upon in the theoretical part of the work.

In all three bodies of work, the panorama was shown to occupy a middle-ground between two conflicting roles of the image: a compromise between serving as a projective surrogate of a light field, and acting as a shared medium for multiple viewers.

6. APPENDICES

6.1 APPENDIX A: EXHIBITIONS *OF LA DISPERSION DU FILS*

La Dispersion Du Fils has been exhibited publically on 10 separate occasions:

- Chronus Art Centre, Shanghai, China, June 2014
- Run Run Shaw Creative Media Centre Grand Opening Festival, Hong Kong, April 2012
- Festival d'Avignon, Avignon, France, July 2011
- "Matière-Lumière", Bethune Cultural Capital of Europe, France, April - June 2011
- STRP Festival, Eindhoven, Netherlands, Nov 2010
- ALIVE Lab Grand Opening, City University Hong Kong, June 2010
- Digital Life, La Pelanda, Rome, Italy, March - April 2010
- Seconde Nature, Fondation Vasarely, Aix-en-Provence, France, July 2009
- Le Volcan Numerique, Le Havre, France, June 2009
- eArts Festival, Shanghai, China, October 2008

6.2 APPENDIX B: OVID'S ACTAEON

Metamorphoses
By Ovid Written 1 A.C.E.

Translated into English verse under the direction of Sir Samuel Garth by John Dryden, Alexander Pope, Joseph Addison, William Congreve and other eminent hands (1717).

Book III:138-252

Actaeon was the first of all his race,
Who griev'd his grandsire in his borrow'd face;
Condemn'd by stern Diana to bemoan
The branching horns, and visage not his own;
To shun his once lov'd dogs, to bound away,
And from their huntsman to become their prey,
And yet consider why the change was wrought,
You'll find it his misfortune, not his fault;
Or, if a fault, it was the fault of chance:
For how can guilt proceed from ignorance?

The Transformation of Actaeon into a Stag

In a fair chace a shady mountain stood,
Well stor'd with game, and mark'd with trails of blood;
Here did the huntsmen, 'till the heat of day,
Pursue the stag, and load themselves with rey:
When thus Actaeon calling to the rest:
"My friends," said he, "our sport is at the best,
The sun is high advanc'd, and downward sheds
His burning beams directly on our heads;
Then by consent abstain from further spoils,
Call off the dogs, and gather up the toils,
And ere to-morrow's sun begins his race,
Take the cool morning to renew the chace."
They all consent, and in a chearful train
The jolly huntsmen, loaden with the slain,
Return in triumph from the sultry plain.

Down in a vale with pine and cypress clad,
Refresh'd with gentle winds, and brown with shade,
The chaste Diana's private haunt, there stood
Full in the centre of the darksome wood
A spacious grotto, all around o'er-grown
With hoary moss, and arch'd with pumice-stone.

From out its rocky clefts the waters flow,
And trickling swell into a lake below.
Nature had ev'ry where so plaid her part,
That ev'ry where she seem'd to vie with art.
Here the bright Goddess, toil'd and chaf'd with heat,
Was wont to bathe her in the cool retreat.

Here did she now with all her train resort,
Panting with heat, and breathless from the sport;
Her armour-bearer laid her bow aside,
Some loos'd her sandals, some her veil unty'd;
Each busy nymph her proper part undrest;
While Crocale, more handy than the rest,
Gather'd her flowing hair, and in a noose
Bound it together, whilst her own hung loose.
Five of the more ignoble sort by turns
Fetch up the water, and unlade the urns.

Now all undrest the shining Goddess stood,
When young Actaeon, wilder'd in the wood,
To the cool grott by his hard fate betray'd,
The fountains fill'd with naked nymphs survey'd.
The frighted virgins shriek'd at the surprize
(The forest echo'd with their piercing cries).
Then in a huddle round their Goddess prest:
She, proudly eminent above the rest,
With blushes glow'd; such blushes as adorn
The ruddy welkin, or the purple morn;
And tho' the crowding nymphs her body hide,
Half backward shrunk, and view'd him from a side.
Surpriz'd, at first she would have snatch'd her bow,
But sees the circling waters round her flow;
These in the hollow of her hand she took,
And dash'd 'em in his face, while thus she spoke:
"Tell, if thou can'st, the wond'rous sight disclos'd,
A Goddess naked to thy view expos'd."

This said, the man begun to disappear
By slow degrees, and ended in a deer.
A rising horn on either brow he wears,
And stretches out his neck, and pricks his ears;
Rough is his skin, with sudden hairs o'er-grown,
His bosom pants with fears before unknown:
Transform'd at length, he flies away in haste,
And wonders why he flies away so fast.
But as by chance, within a neighb'ring brook,

He saw his branching horns and alter'd look.
Wretched Actaeon! in a doleful tone
He try'd to speak, but only gave a groan;
And as he wept, within the watry glass
He saw the big round drops, with silent pace,
Run trickling down a savage hairy face.
What should he do? Or seek his old abodes,
Or herd among the deer, and sculk in woods!
Here shame dissuades him, there his fear prevails,
And each by turns his aking heart assails.

As he thus ponders, he behind him spies
His op'ning hounds, and now he hears their cries:
A gen'rous pack, or to maintain the chace,
Or snuff the vapour from the scented grass.

He bounded off with fear, and swiftly ran
O'er craggy mountains, and the flow'ry plain;
Through brakes and thickets forc'd his way, and flew
Through many a ring, where once he did pursue.
In vain he oft endeavour'd to proclaim
His new misfortune, and to tell his name;
Nor voice nor words the brutal tongue supplies;
From shouting men, and horns, and dogs he flies,
Deafen'd and stunn'd with their promiscuous cries.
When now the fleetest of the pack, that prest
Close at his heels, and sprung before the rest,
Had fasten'd on him, straight another pair,
Hung on his wounded haunch, and held him there,
'Till all the pack came up, and ev'ry hound
Tore the sad huntsman grov'ling on the ground,
Who now appear'd but one continu'd wound.
With dropping tears his bitter fate he moans,
And fills the mountain with his dying groans.
His servants with a piteous look he spies,
And turns about his supplicating eyes.
His servants, ignorant of what had chanc'd,
With eager haste and joyful shouts advanc'd,
And call'd their lord Actaeon to the game.
He shook his head in answer to the name;
He heard, but wish'd he had indeed been gone,
Or only to have stood a looker-on.
But to his grief he finds himself too near,
And feels his rav'nous dogs with fury tear
Their wretched master panting in a deer.

6.3 APPENDIX C: VØSPAZÀR TEXT

YEAR 1999 A.D

1. The artists of LFK declare that they are dogs.
2. They affirm that they are descended from Melampus, Ichnobates, Pamphagos, Dorceus, and Oribasos, Nebrophonos, Theron, Lælaps, Pterelas and from all those who have eaten the hunter in the prey, the master in the beast.
3. They claim to be direct descendents from this pack, the one which, having thus driven the negating, synergetic, and initiatory action right to its end, was the inventor of the first deed of the *Gesamtkunstwerk*.
4. They recall that, making their meal of the man-animal, of the Being-non-Being in which the image of divine nudity was enclosed, the dogs of Gargaphie had also devoured and thus had had to assimilate a mystical destiny. For, those dogs carried in their stomachs the transformed material of an order, of a challenge, of a justly »cynical« and divine authorization bestowed upon the hunter become tawny and, finally, eaten: that of showing what cannot be told, that is to say, the mystical itself, if Wittgenstein is to be believed.
5. It is on the basis of their particular canine ancestry that the members of LFK found and justify their collective interest in the arts in general, and in filmmaking in particular, insofar as it be essentially the art of silence.
6. Until that time a poor group with neither artistic place nor aesthetic morale, a simple and errant band already practicing negativity but in a purely instinctive way, the artists of LFK henceforth claim their autonomy as a cynical people on the territories of film. Following no master, faithful to their only ancestors, they conceive of a cyno-gestural order, a kind of cynema, that they uselessly call Vøspazà and whose primary characteristic feature is the complete incomprehensibility of its language [mysterious to itself].

7. The bit by bit, element by element expulsion and progressive evacuation of the faeco-philosophical and excreto-mystical material of devoured myth will one day constitute the essential deed of the repertoire of the Vøspazàrian film-makers [cynéaste]. This will be intellectuo-convulsive, above all.
8. Given how complicated and fastidious it seems to be to want to reconstitute the cervidae body of Actæon as the product of extractions from a multitude of stomachs, how impossible it would be in practice to bring him back to life and how, despite all resurrection, he would nevertheless still be incapable of telling what he saw as a man, the artist-dogs of LFK renounce straightaway both coherence and narration in their work. As well, it is not to extracts that they attribute an autonomous value, but to extraction itself. The value of the pure deed, for which they seek no end other than itself.

Jean Michel Bruyère (2003, p. 390).

6.4 APPENDIX D: VøSPAZÀR CYCLE EXHIBITIONS AND WORKS

Listed here are the various installations, exhibitions, films and theatre works performed by Bruyère and the LFKs as part of the VøSPAZÀR cycle of works:

I. ACTION (A), INSTALLATION (I), EXHIBITION (EP), MEDIA ARTS (AM), IMMERSIVE THEATRE (TI)

La Dispersion Du Fils (AM)(I)
Shanghai, Chronus Art Centre, 2014
Hong Kong, Run Run Shaw Creative Media, 2012
Avignon, Festival d'Avignon, 2011
Bethune, "Matière-Lumière", 2011
Eindhoven, STRP Festival, 2010
Hong Kong, ALIVE City University, 2010
Rome, Digital Life, La Pelanda, 2010
Aix-en-Provence, Seconde Nature, Fondation Vasarely, 2009
Le Havre, Le Volcan Numerique, 2009
Shanghai, eArts Festival, 2008

CaMg[CO3]2 (AM)(I)
Marseille, la Friche la Belle de Mai, 2013
Le Havre, Immersion, Le Volcan, 2008
Mulhouse, Trans(e), La Filature, 2008
Karlsruhe, Panorama Festival, ZKM 2008
Berlin, Vom Funken Zum Pixel, Martin Gropius Bau, 2007-2008
Aix-en-Provence, Festival Seconde Nature, 2007

À l'Enseigne des Vrais Chiens (A)(I)
Berlin, Haus der Kulturen der Welt, St. Elisabeth-Kirche, 2007

7 Fuites du Paysage (I)(EP)
Aix-en-Provence, Arborescence, Cézanne 2006, Chapelle des Pénitents Blancs, 2006

Une Défaite (EP)
Aix-en-Provence, Arborescence, Cézanne 2006, Galerie 200RD10, Vauvenargues, 2006

Vi Summa Vocis (A)(I)
Munich, Muffathalle, 2006

L'Insulte Faite au Paysage (TI)(I)
Avignon, Festival d'Avignon, Église des Célestins, 2005

Sui in Res (A)(I)(EP)
Avignon, Festival d'Avignon, Chapelle Saint-Charles, 2004

One Deer Nine Dogs (A)(I)(EP)
Marseille, Le Merlan Scène Nationale, 2004

Si Poteris Narrare, Licet (AM)
Le Havre, Immersion, le Volcan Scène Nationale, 2008
Berlin, Vom Funken Zum Pixel, Martin Gropius Bau, 2007-2008
Aix-en-Provence, Arborescence, Cézanne 2006, Fondation Vasarely, 2006
Shanghai, MOMA, 2006
Lille 2004
Rotterdam, Deaf, 2004
Belfort, CICV, 2003
Athènes, Medi@terra Festival, 2003
Karlsruhe, Future Cinema, ZKM, 2003
Créteil, Festival Exit, Maison des Arts, 2002
Maubeuge, Festival Via, La Luna, 2002

Une Brutalité Pastorale (TI)(I)(EP)
Belfort, Interférences, Forêts de Vézelois, 2000

Elements of a Naked Chase (A)(I)
Marseille, Le Merlan Scène Nationale, 1999

Ecran-Carne (A)(AM)
Belfort, Nuits Savoureuses, 1999

II. CINEMA (C), IMMERSIVE CINEMA (CI), VIDEO (V), TEXT (T)

La Dispersion Du Fils **(CI)**, 2009
CaMg(CO3)2 **(CI)**, 2007
L'Insulte Faite Au Paysage **(C)**, 2005
La Tragédie d'Actéon **(V)**, 2005
Acedia **(T)**, 2005
La Constellation de l'Âne **(V)**, 2005
7 Fuites du Paysage **(V)**, 2004
Barking Forefathers **(V)**, 2004
Si Poteris Narrare, Licet **(CI)**, 2002
Première Sphère des Mouvements Fixes **(V)**, 2001
Natural Elements of Vospàzà **(V)**, 2000
Ecran Carne **(V)**, 2000
Elements of a Naked Chase **(C)**, 1999
Quelques Cyclopes **(V),** 1999

See *www.epidemic.net* and *www.lfks.net* for more information concerning LFKs.

6.5 APPENDIX E: PROJECTOR BASICS

The *frustum* of a projector, the pyramidal volume of light it produces, is a useful conceptual tool when working with projectors. A projector's frustum is a rectangular oblique pyramid whose apex, the *centre-of-projection,* is the focal point of the lens, and whose angles are determined by three properties of a projector: the aspect-ratio of the projector's image panel, the focal length of the projector's lens, and the projector's lens shift. Aspect-ratio determines the ratio of the width and height of the frustum. At the time the initial design of the system was undertaken, most digital projectors enjoyed an aspect ratio of 4:3, while at the time or writing, 16:9 and 16:10 aspect ratios have become the norm. Focal length determines the *throw* of the projector, which is the ratio of the width of the image formed over the distance of projection. Projector lenses are often described in terms of throw, rather than field-of-view or focal length, and typical values range from 0.75:1 for a 16mm ultra-wide angle "short-throw" lens to 6.5:1 for a 136mm telephoto "long-throw" lens.

Lens-shift is the displacement, parallel to the *optical plane* and perpendicular to the *optical axis*, of the centre of the image away from the optical centre of the lens.[226] This has the effect of translating the projected image a proportionally equivalent amount. For example, shifting the image-panel upward by half its height upward displaces the projected image by half its height, downward. As such, lens-shift is measured in percentages, a 50% vertical shift indicating, for example, that the image-panel, and

[226] Lens-shift is often achieved by mechanically displacing the lens, rather than the image panel, hence the name. The results are identical.

therefore the projected image, have been displaced upwards by 50% of their height. Note that lens-shift results in an asymmetrical frustum, as the apex is now displaced from the centre of the image.

Most projectors are fixed, for convenience, with a vertical lens shift of around +/-50%, so that the top or bottom of the projected image rests in line with the centre of the lens. This allows a projector to be placed in line with the top of a projection screen, where it can project over the heads of the audience, and still produce a rectangular image. More expensive projectors allow the adjustment of the lens-shift, not only in the vertical, with typical ranges such as +/- 110%, but also in the horizontal, albeit with a less extensive range.

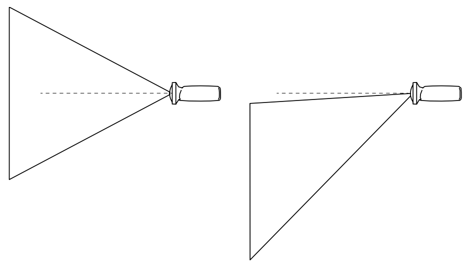
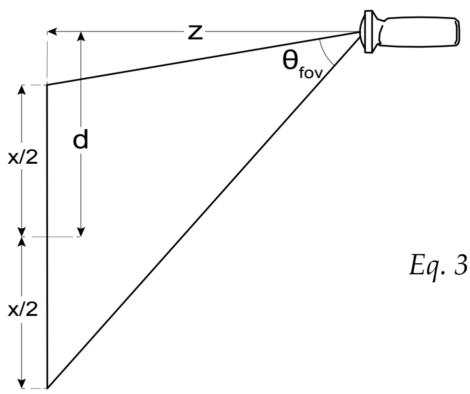


Figure 140 - Lens Shift.
Left: Zero lens-shift. The image-panel and frustum are centred on the optical axis.
Right: ~55% vertical lens-shift, a typical configuration for projectors. The image-panel is shifted upwards, resulting in the projected image shifting down.

It is sometimes useful to describe a projector in terms of its horizontal and vertical fields-of-view, which are the angles formed at the pivot of the frustum. However, a little trigonometry reveals that these angles vary with lens-shift, which explains why throw is a more useful parameter than field-of-view.



If throw t = x / z and lens shift s = d / x, then field of view θ_{fov} varies with both s and t:

$$\theta_{fov} = 2\left(\tan^{-1}\left(\frac{t}{2} + s.t\right) - \tan^{-1}(s.t)\right)$$

Eq. 30 Field-of-view as a function of throw and lens shift.

6.6 APPENDIX F: OMNISTEREO FRUSTUM OVERLAP

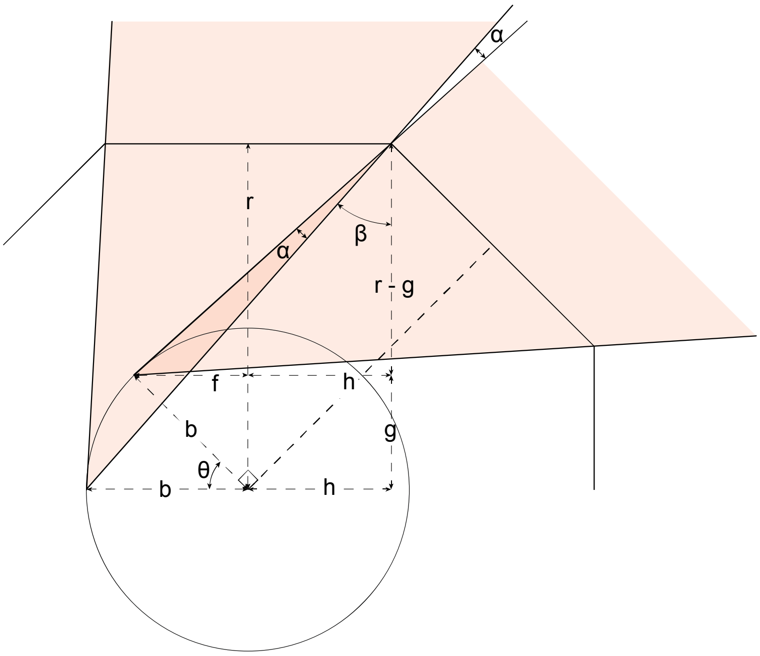


Figure 141 - Measure of overlap between neighbouring frusta.

If $f = b\cos\theta,\ g = b\sin\theta,\ h = r\tan\frac{\theta}{2}\ ,\beta = \tan^{-1}\left(\frac{b+h}{r}\right)$, and $\theta = \frac{2\pi}{n}$, where n is the number of slices, r the (inner) screen radius and b half the stereo base-line, then

$$\tan(\alpha + \beta) = \frac{f+h}{r-g}$$

Substitution yields:

$$\alpha = \tan^{-1}\left(\frac{b\cos\theta + r\tan\frac{\theta}{2}}{r - b\sin\theta}\right) - \tan^{-1}\left(\frac{b + r\tan\frac{\theta}{2}}{r}\right)$$

Which, with $w = \frac{b}{r}$ and $= \frac{2\pi}{n}$, gives:

$$\alpha = \tan^{-1}\left(\frac{w\cos\theta + \tan\frac{\theta}{2}}{1 - \mathrm{w}\sin\theta}\right) - \tan^{-1}\left(\mathrm{w} + \tan\frac{\theta}{2}\right)$$

6.7 APPENDIX G: ICINEMA SDK

Listed here are the various modules that together constitute the iCinema SDK. Most take the form of Virtools plug-ins, extending the functionality of Virtools to better support, and accelerate the development of AVIE applications.

icBase

Base infrastructure including logging, debugging, maths, graphics and file utilities as well as C++ base-classes for smart parameters, smart objects, pointers & containers. Also includes XML and optimised file-streaming tools.

An oft used feature of icBase is the icInterpolator system, which allows any numerical value to be set over time with a single function call. Experience developing immersive real-time systems has led to the observation that it is very uncommon that properties of the virtual world are desired to change state instantaneously, but rather should change over time. The icInterpolator system allows the developer to set any variable in motion towards some desired value, with a single function call.

icAvieBase / icDomeBase

A single *Virtools* composition that unifies all the iCinema SDK features relevant or required for an AVIE experience. icAVIEBase provides a 3D graphical user interface that permits real-time adjustment and monitoring of all important iCinema SDK variables.

This composition is designed to host 'client' compositions, and in doing so, relieves the developer of the client composition from re-implementing these features. By separating the AVIE specific code from the application specific code (between system and content), development of AVIE experiences is accelerated. In fact, certain types of experiences can be created without recourse to any of the iCinema SDK, and yet, by hosting them in icAVIEBase, experienced in the AVIE.

In addition, icAVIEBase allows developers to create AVIE applications on standard desktop machines. For example, it provides tools for simulating the cluster on a single computer, or simulating AVIE input devices with mouse or keyboard.

icAVIEBase is also completely interchangeable with its counterpart icDomeBase, which plays the same role for the iCinema iDome platform. As such, an application intended for the AVIE can be viewed on the iDome without change, and vice versa.

icAI

icAI is a framework for animating and controlling autonomous virtual characters. The module contains a hierarchical behaviour scheduling system, an animation system including dynamic foot placement, grasping, simulation of attentional gaze with head, eye and body movements, rag-doll physics and behaviours for navigation and following moving targets. It also contains interfaces for communicating with external AI systems (eg. Prolog programs).

icScenarioManager / icEntityManager

A framework for reasoning about "spatial narratives". Spatial events - collisions and separations, accelerations and trajectories, approaches and departures, entrances and exits - are automatically detected, classified, and then broadcast to subscribers. Events can be configured to trigger automatic responses (sounds, graphics), but more importantly, autonomous agents can reason about and react to these events. As human and virtual agents are equally modelled as narrative agents, narrative events can be detected in both the virtual and real world, or between the two. This provides a framework for reasoning about and constructing 'co-evolutionary narratives' as a series of events and interactions unfolding in the virtual and real and across the divide.

In addition, a framework for handling large groups of homogeneous entities and agents is provided. This framework allows for simple implementation of group dynamics with emergent properties, such as flocking, for example. The system allows hierarchical nesting of groups by treating groups as entities themselves. This permits agents to reason about or react to groups as a whole, or to their individual constituents. In addition, the system provides a mechanism for off-loading computation to the GPU, so that real-time simulation of systems containing millions of entities is feasible. As such, this system provides a unified scheme for modelling collections of entities of arbitrary complexity or autonomy, from groups of intelligent agents to particle systems.

icField / icFieldManager / icNoiseField

A unified framework for modelling scalar and vector fields. The icField base class provides a unified interface for modelling, managing and animating various field-like phenomena such as wind, gravity, water turbulence, magnetic fields. The class also provides GUI menus, visual debugging and automatic management of GPU shader parameters.

icField includes implementations of commonly used noise fields (Perlin and Simplex based fractal Brownian motion, ridged fractal and turbulence) in one, two, three and four dimensions. Matching CPU and GPU versions of these noise fields are provided.

icAvieConfig

A stand-alone application for calibrating projection distortion and blending of multiple projectors in order to create a single seamless image. icAvieConfig is described in more detail in Section 3.9.5.

icAVIETrack

icAVIETrack provides an interface between the *icScenarioManager* and the *Immersitrack* tracking system (Sridhar, 2012). See Section 3.12.1 for more details.

icClikapad

Interface to the ClikaPad radio-frequency input device.[227]

icCluster / icDistributed / icMessageManager

The icCluster, icDistributed and icMessage modules provide tools for distributing a *Virtools* application across a cluster of machines. They do so by extending the *Virtools* native clustering mechanisms with a number of new features:

- Distributed parameter and object base-classes that automate the distribution of data and object state
- A distributed messaging system for effecting remote-procedural calls

[227] See *www.clikapad.com*

- Compression and decompression of data to minimise data bandwidth
- Automatic detection of unequal calls to random number functions, which can result in the states of different computers diverging
- Simulation of cluster on a single machine for debugging purposes
- Automatic distribution of icGUI

icComm

An inter-computer messaging system based on TCP/IP. In contrast to the UDP protocol used by icCluster, TCP/IP allows for the transport of arbitrarily large amounts of data, and verifies that the data arrives intact, at the expense of speed and latency.

icExample

A template for rapidly creating new iCinema SDK modules.

icGUIVT

The icGUI system is a 3D Graphical User Interface (GUI) includes a simple menu system for exposing system and application parameters so they can be easily manipulated during the running of a program.

The GUI is cluster-compatible, and viewable from both the master computer and within the AVIE itself. As the GUI is itself a 3D structure, it can be manipulated using any of the AVIE input devices, from mouse to a 3D pointer to the full-body camera-based tracking system.

The icGUI system provides methods for transforming any collection of virtual objects into a physical interface. First, a virtual object is opened to physical manipulation of its position, orientation or size using any of the AVIE's input devices. This can be in a manner that simulates real-life physics (simulation of mass, inertia, collisions, friction, spring forces or joints with constraints), or according to arbitrary user-defined rules and constraints. Then, specific qualities of this virtual object are linked to some arbitrary variable of the application or system, allowing for the alteration or visualisation of this variable through the physical movement of the object. For example, the size of a sphere might be linked to the sound engine's volume, or the orientation of a cube linked to the velocity of the AVIE through space. The power of this system lies in the ability to embed a GUI seamlessly within the virtual world, maintaining the physical activity, consistency and plausibility so important to presence.

icIntersenseVT

An interface for Intersense family of tracking devices, such as the Inertiacube 3-DOF orientation sensor used in the AVIE wand devices (see Section 3.12).

icLightManager

A real-time shader-based shadow system providing soft shadows cast from point, directional and spot lights.

iCME

A multi-threaded video engine providing high-bandwidth streaming of movies encoded with a custom codec based on the DXTn (S3tc) texture format. As this texture format is decompressed entirely on the GPU and not on the CPU, the limitations in bandwidth caused by CPU decompression, or the movement of data from RAM to VRAM, is essentially removed. The engine supports ultra-high resolution videos (e.g. 8K x 4k) by subdividing frames into tiles and hierarchical resolutions, and allows distribution of movie data across multiple disks for maximum performance. It is able to stream just those portions of the video frame that are visible at any moment, both from disk to RAM, and from RAM to VRAM. It also provides multi-threaded playback of very high numbers of videos simultaneously. Lastly, iCME includes stand-alone tools for the compression and previewing of content.

iCME was used to great effect in the iCinema project Conversations (Del Favero, Gibson, Shaw 2004), where omnistereoscopic 8k x 4k resolution spherical videos were viewed using a head mounted display.

icMPEG2

A high-performance mpeg2 video engine, icMPEG2 is designed to play both high-resolution videos and large numbers of videos simultaneously. As videos are streamed to the GPU as textures, they can be used in any manner within a 3D scene;

mapped to surfaces, or as lights, or as abstract data, for example. The system also provides GPU accelerated colour correction and colour space conversion.

The system supports arbitrary video resolutions, limited only by the processing power and maximum texture width of the graphics card. When videos are mapped to entities in motion, the system automatically calculates which videos are visible and which are not on any given computer, and streams only the required data to the graphics card. This is particular useful for the AVIE, where each computer in the cluster is responsible for just one sixth of the panoramic screen. The engine provides real-time performance metrics and debugging data, as well as stand-alone benchmarking tools, for assessing where any bottlenecks may lie as video data makes its journey from disk to screen. Potential bottlenecks lie in disk access, CPU decoding, RAM to VRAM transfer, or rendering itself, any of which may present themselves as bottleneck, depending on the context.

A large amount of time was devoted to solving problems of parallelisation and synchronisation across multiple computers. As the multi-threaded design of the movie engine adds an element of indeterminism to the cluster state - for one computer may load data faster than another - and different videos are visible at different times on different computers, great care must be taken to avoid any discrepancies arising between computers. The synchronisation system allows for fine control over skipping of frames, synchronisation with soundtracks and frame-accurate seeking.

A equally large amount of time was devoted to ensuring operations of videos such as loading, closing, pausing, playing, seeking or looping do not disrupt the smoothness of motion of the system as a whole. Any jumps or jitters in the AVIE frame-rate are highly visible and disrupt sense of presence. Where possible, the video engine fulfils its tasks asynchronously in worker threads, and where not, pools of resources (threads, VRAM textures, memory buffers) are created during initialisation and reused, so as to avoid disruptive creation or destruction of such resources during runtime.

icOSCVT

An Open-Sound-Control plugin for *Virtools*. OSC is a commonly used protocol for inter-computer or inter-process communication (Wright, 2005).

icShaderManager

Extends the *Virtools* shader system with a number of important features, including off-line multi-threaded compilation, loading of pre-compiled shaders, automatic binding of shader parameters to the icGUI, management of static shader parameters, saving and loading shaders from external files.

icSoundEngine

A 3D sound engine capable of driving any configuration of speakers up to 12.1. Described in more detail in Section 3.10.5.

icSpeechVT

An interface between *Virtools* and Microsoft Speech Recognition API.

icToolkitVT

The icToolkitVT module is a collection of miscellaneous functions and classes that do not belong in any of the other modules. This includes such things as a large library of helper classes and functions for rapidly creating geometry, textures or materials, for sending debug information to log files, the console or displaying it graphically in the AVIE and a system for reading values from configuration files, among many other miscellanies.

icVoxCogVT

A speech recognition module based on the Dragon speech system

icWarpVT

The icWarp system provides the functions required for the rendering of images for the AVIE and iDome systems. It includes tools for distortion correction and projection blending, and the real-time production of fisheye (for the iDome) and omnistereoscopic (for the AVIE) images. icWarp reads the XML configuration files created by icAVIEConfig calibration system.

The icWarp system provides a structure of classes for modelling projection systems, including representations of projectors, render-nodes, render-windows and rendering algorithms (render "views"). In abstracting the notion of a projection system, applications can be migrated from one VR configuration to another without modification.

icWarp is completely integrated with the icCluster system. It supports synchronised rendering across a cluster of computers, or the creation of virtual clusters on a single computer. A virtual cluster, where multiple instances of the application run on the same computer, is typically used for development and debugging, but it can also be considered a naive approach to parallelisation, and in some cases might be used to provide performance benefits.

As each render-node is responsible for calculating just a subsection of the panorama, icWarp provides an automatic culling system that accelerates rendering by determining local visibility for each render-node.

icWarp provides real-time control of all rendering parameters through the icGUI systems and the API. It provides a variety of development tools to accelerate development of AVIE applications, such as the ability to render the view of any render-node on the master computer, or display debug render data and structure within the AVIE. It also provides tools for monitoring frame-rate and frame-jitter, and control over synchronisation methods.

The system provides support for stereo objects - entities, meshes, materials or textures that have different appearances for the left and right eye. This can be used, for example, to embed a stereo photograph within a 3D scene.

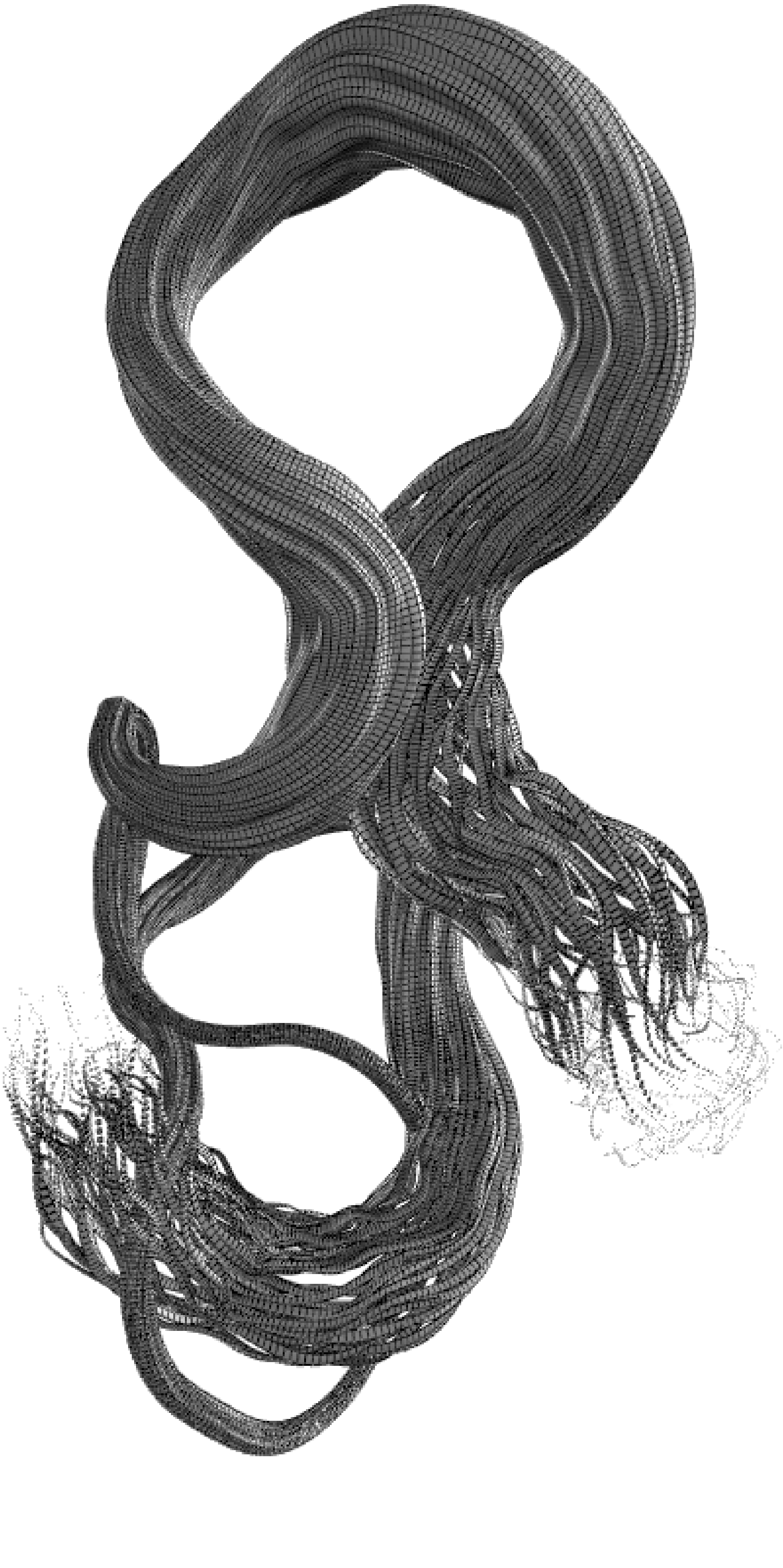
The system also provides a system for managing visibility of render-sets; the linking of render-passes with sets of objects.

icWarp also provides a post-processing framework. Rendered scenes are not sent directly to the projectors, but buffered in textures, where they can be further manipulated in what is called a 'post-processing' pass. Effects include gamma and colour correction, blur and motion blur, glow and bloom and (de)saturation.

Finally, icWarp provides a system for capturing images, either as complete panorama, or individual images from each render-node.

icWebVT

icWebVT provides an interface to the world-wide web from within a *Virtools* application. It provides asynchronous non-blocking methods for requesting data (e.g. from *Google* or *FlickR*) and then subsequent reception of results. It also provides mechanisms for embedding web content within the virtual world, by allowing the streaming of HTML, images, video and Flash content onto arbitrary 3D surfaces.



7. BIBLIOGRAPHY

Adams, K. R. (1972). Perspective and the Viewpoint. *Leonardo,* 5(3), 209-217.

Adelson, E. H., & Bergen, J. R. (1991). The Plenoptic Function and the Elements of Early Vision. *Computational Models of Visual Processing,* 91(1), 3-20.

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from near-Optimal Bimodal Integration. *Current Biology,* 14(3), 257-262.

Allard, J., Gouranton, V., Lamarque, G., Melin, E., & Raffin, B. (2003). Softgenlock: Active Stereo and Genlock for PC Cluster. *Proceedings of the Workshop on Virtual Environments 2003*, 255-260.

Allard, J., Gouranton, V., Lecointre, L., Melin, E., & Raffin, B. (2002). Net Juggler: Running VR Juggler with Multiple Displays on a Commodity Component Cluster. *Proceedings of IEEE Virtual Reality 2002*, 273-274.

Allard, J., Lesage, J.-D., & Raffin, B. (2010). Modularity for Large Virtual Reality Applications. *Presence: Teleoperators and Virtual Environments,* 19(2), 142-161.

Allard, J., Ménier, C., Boyer, E., & Raffin, B. (2005). Running Large VR Applications on a PC Cluster: The Flow VR Experience. *Proceedings of the 11th Eurographics Conference on Virtual Environments*, 59-68.

Anderton, J. (1895). *Method by Which Pictures Projected Upon Screens by Magic Lanterns Are Seen in Relief*, U.S.A., Patent No. 542,321.

Antani, L., Chandak, A., Taylor, M., & Manocha, D. (2012). Direct-to-Indirect Acoustic Radiance Transfer. *IEEE Visualization and Computer Graphics,* 18(2), 261-269.

Apostol, T. M., & Mnatsakanian, M. A. (2007). Unwrapping Curves from Cylinders and Cones. *American Mathematical Monthly,* 114(5), 388-416.

Archer, B. (1995). The Nature of Research. *Co-Design Journal,* 2(11).

Ardouin, J., Lécuyer, A., Marchal, M., & Marchand, E. (2014). Stereoscopic Rendering of Virtual Environments with Wide Field-of-Views up to 360°. *Proceedings of IEEE Virtual Reality 2014*, Minneapolis, MN, 3-8.

Aymerich-Franch, L. (2010). Presence and Emotions in Playing a Group Game in a Virtual Environment: The Influence of Body Participation. *Cyberpsychology, Behavior, and Social Networking,* 13(6), 649-654.

Aymerich-Franch, L. (2012). Can We Identify with a Block? Identification with Non-Anthropomorphic Avatars in Virtual Reality Games. *International Society for Presence Research Annual Conference*, Philadelphia, Pennsylvania, USA.

Baalman, M. (2007). On Wave Field Synthesis and Electro-Acoustic Music: State of the Art 2007. *Proceedings of the International Computer Music Conference 2007*.

Balzer, R. (1998). *Peepshows: A Visual History*. New York: Harry N. Abrams.

Banks, M. S., Read, J. C., Allison, R. S., & Watt, S. J. (2012). Stereoscopy and the Human Visual System. *SMPTE Motion Imaging Journal,* 121(4), 24-43.

Banks, M. S., Rose, H. F., Vishwanath, D., & Girshick, A. R. (2005). Where Should You Sit to Watch a Movie? *Proc. SPIE 5666, Human Vision and Electronic Imaging X*, 316-325.

Barco. (2013). Barco's Complete Range of Stereoscopic Technologies. Accessed September 2013 at http://www.barco.com/projection_systems/downloads/stereoscopic_technologies_overview.pdf

Barker, R. (1796). Specification of the Patent Granted to Mr. Robert Barker ... Called by Him La Nature À Coup D'œil. *The Repertory of Arts and Manufactures: Consisting of Original Communications, Specifications of Patent Inventions, and Selections of Useful Practical Papers From The Transactions of the Philosophical Societies of all Nations.,* 4, 165-167.

Baron, A. B. (1897). *Système D'appareil Pour Projections Circulaires Animées Dit Cinématorama*, France, Patent No. 261,650.
Baron, A. B. (1899). *Système D'appareil Pour Projections Panoramiques Circulaires Animées, En Couleurs Et Parlantes, Dit Cinématorama Parlant.*, France, Patent No. 294,384.
Barthes, R. (1968). L'effet De Réel. *Communications,* 11(1), 84-89.
Bassili, J. N. (1976). Temporal and Spatial Contingencies in the Perception of Social Events. *Journal of Personality and Social Psychology,* 33(6), 680.
Battisti, E., Battisti, G. S., & Fontana, G. (1984). *Le Macchine Cifrate Di Giovanni Fontana: Con La Riproduzione Del Cod. Icon. 242 Della Bayerische Staatsbibliothek Di Monaco Di Baviera E La Decrittazione Di Esso E Del Cod. Lat. Nouv. Acq. 635 Della Bibliothèque Nationale Di Parigi*: Arcadia Edizioni.
Bauer, B. (1962). Phasor Analysis of Some Stereophonic Phenomena. *IRE Transactions on Audio*(1), 18-21.
Bazin, A. (1967). *What Is Cinema?* (H. Gray, Trans. Vol. 1). Berkeley: University of California Press.
Bellefeuille, A., & Faubert, J. (1998). Independence of Contour and Biological-Motion Cues for Motion-Defined Animal Shapes. *PERCEPTION-LONDON-,* 27, 225-236.
Bennet, J. C., Barker, K., & Edeko, F. O. (1985). A New Approach to the Assessment of Stereophonic Sound System Performance. *Journal of the Audio Engineering Society,* 33(May), 314-321.
Bennett, J. (2008). *T_Visionarium: A User's Guide*. Sydney: NewSouth Publishing.
Berkhout, A. J. (1988). *A Holographic Approach to Acoustic Control*. New York: Audio Engineering Society.
Berkhout, A. J., De Vries, D., & Vogel, P. (1993). Acoustic Control by Wave Field Synthesis. *The Journal of the Acoustical Society of America,* 93(May), 2764–2778.
Bernfeld, B. (1973). Attempts for Better Understanding of the Directional Stereophonic Listening Mechanism*. Audio Engineering Society Convention 44*, Rotterdam, 308.
Berry, D. S., Misovich, S. J., Kean, K. J., & Baron, R. M. (1992). Effects of Disruption of Structure and Motion on Perceptions of Social Causality. *Personality and Social Psychology Bulletin,* 18(2), 237-244.
Bidet-Caulet, A., Voisin, J., Bertrand, O., & Fonlupt, P. (2005). Listening to a Walking Human Activates the Temporal Biological Motion Area. *Neuroimage,* 28(1), 132-139.
Bierbaum, A., Hartling, P., Morillo, P., & Cruz-Neira, C. (2005). Implementing Immersive Clustering with VR Juggler. *Computational Science and Its Applications–ICCSA 2005* (pp. 1119-1128): Springer.
Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A., & Cruz-Neira, C. (2001). VR Juggler: A Virtual Platform for Virtual Reality Application Development*. Proceedings of IEEE Virtual Reality 2001*, 89-96.
Biocca, F., Kim, J., & Choi, Y. (2001). Visual Touch in Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions. *Presence: Teleoperators and Virtual Environments,* 10(3), 247-265.
Bishop, R. L. (1975). There Is More Than One Way to Frame a Curve. *American Mathematical Monthly,* 82(3), 246-251.
Blake, R., & Shiffrar, M. (2007). Perception of Human Motion. *Annu. Rev. Psychol.,* 58, 47-73.
Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press.
Blauert, J., & Rabenstein, R. (2012). Providing Surround Sound with Loudspeakers: A Synopsis of Current Methods. *Archives of Acoustics,* 37(1), 5-18.
Bletzinger, K.-U. (1998). *Form Finding and Optimization of Membranes and Minimal Surfaces*: Institut für Baustatik.

Blinn, J. F. (1977). Models of Light Reflection for Computer Synthesized Pictures. *SIGGRAPH Comput. Graph.,* 11(2), 192-198.

Blinn, J. F. (1978). Simulation of Wrinkled Surfaces*. ACM SIGGRAPH Computer Graphics*, 12(3), 286-292.

Blinn, J. F., & Newell, M. E. (1976). Texture and Reflection in Computer Generated Images. *Communications of the ACM,* 19(10), 542-547.

Blumlein, A. D. (1958). Improvements in and Relating to Sound-Transmission, Sound-Recording and Sound-Reproducing Systems (British Patent Specification 394,325). *Journal of the Audio Engineering Society,* 6(2), 91-130.

Bolter, J. D., & Grusin, R. A. (1999). *Remediation : Understanding New Media*. Cambridge, Mass.: MIT Press.

Boone, M. M., Verheijen, E. N., & Van Tol, P. F. (1995). Spatial Sound-Field Reproduction by Wave-Field Synthesis. *Journal of the Audio Engineering Society,* 43(12), 1003-1012.

Bourke, P. (2004). Edge Blending Using Commodity Projectors. Accessed July 21 2013 at http://paulbourke.net/texture_colour/edgeblend/

Braund, M. J. (2008). The Structures of Perception: An Ecological Perspective. *Kritike,* 2, 123-144.

Bronkhorst, A., & Houtgast, T. (1999). Auditory Distance Perception in Rooms. *Nature,* 397(6719), 517-520.

Brown, N., Del Favero, D., McGinity, M., Shaw, J., & Weibel, P. (2008). *TVisionarium II* [AVIE-based artwork].

Brown, N. O. (1972). Metamorphoses II: Actaeon. *The American Poetry Review,* 1(1), 38-40.

Brown, R. (1828). A Brief Account of Microscopical Observations Made in the Months of June, July and August 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies. *The Philosophical Magazine, or Annals of Chemistry, Mathematics, Astronomy, Natural History and General Science,* 4(21), 161-173.

Brungart, D. S., & Scott, K. R. (2001). The Effects of Production and Presentation Level on the Auditory Distance Perception of Speech. *The Journal of the Acoustical Society of America,* 110(1), 425-440.

Bruno, G., & Memmo, P. E. (1964). *Giordano Bruno's the Heroic Frenzies. A Translation with Introduction and Notes by Paul Eugene Memmo*: Chapel Hill; Valencia printed.

Bruns, G. (2011). *On Ceasing to Be Human*: Stanford University Press.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*: University of California Pr.

Bruyère, J. M. (2003). Si Poteris Narrare, Licet. (S. Clift, Trans.) In J. Shaw & P. Weibel (Eds.), *Future Cinema: The Cinematic Imaginary after Film*: MIT Press.

Buehner, M. J. (2012). Understanding the Past, Predicting the Future Causation, Not Intentional Action, Is the Root of Temporal Binding. *Psychological science,* 23(12), 1490-1497.

Buehner, M. J., & Humphreys, G. R. (2009). Causal Binding of Actions to Their Effects. *Psychological Science,* 20(10), 1221-1228.

Bues, M., Blach, R., Stegmaier, S., Häfner, U., Hoffmann, H., & Haselberger, F. (2001). Towards a Scalable High Performance Application Platform for Immersive Virtual Environments*. Proceedings of the 7th Eurographics conference on Virtual Environments & 5th Immersive Projection Technology*, 165-174.

Burr, D., & Thompson, P. (2011). Motion Psychophysics: 1985–2010. *Vision research,* 51(13), 1431-1456.

Burtnyk, N., & Wein, M. (1976). Interactive Skeleton Techniques for Enhancing Motion Dynamics in Key Frame Animation. *Communications of the ACM,* 19(10), 564-569.

Byers, S. (2006). Life as" Self-Motion": Descartes and" the Aristotelians" on the Soul as the Life of the Body. *The Review of Metaphysics*, 723-755.
Candy, L. (2006). Practice Based Research: A Guide. *Report from Creativity and Cognition Studios, University of Technology, Sydney. Retrieved January,* 19, 2010.
Catmull, E., & Rom, R. (1974). A Class of Local Interpolating Splines. *Computer aided geometric design,* 74, 317-326.
Cavazza, M., Lugrin, J.-L., & Buehner, M. (2007). Causal Perception in Virtual Reality and Its Implications for Presence Factors. *Presence: Teleoperators and Virtual Environments,* 16(6), 623-642.
Cavell, S. (1979). *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*: Oxford University Press.
Cecchi, S., Romoli, L., Peretti, P., & Piazza, F. (2012). Low-Complexity Implementation of a Real-Time Decorrelation Algorithm for Stereophonic Acoustic Echo Cancellation. *Signal Processing,* 92(11), 2668-2675.
Chaldecott, J. (1953). The Zograscope or Optical Diagonal Machine. *Annals of Science,* 9(4), 315-322.
Charbonneau, G., Véronneau, M., Boudrias-Fournier, C., Lepore, F., & Collignon, O. (2013). The Ventriloquist in Periphery: Impact of Eccentricity-Related Reliability on Audio-Visual Localization. *Journal of vision,* 13(12), 20.
Scientific American (1896). Chase's Electric Cyclorama (Vol. 74, pp. 120-121).
Chase, C. A. (1895a). *Reproducing Device*, U.S.A., Patent No. 545,422.
Chase, C. A. (1895b). *Stereopticon Panorama Machine*, U.S.A., Patent No. 545,423.
Chase, C. A. (1895c). *Stereopticon Panorama Machine (II)*, U.S.A., Patent No. 545,452.
Western Electrician (1895). Chase Electric Cyclorama (Vol. 17, pp. 235-237).
Chen, H., Chen, Y., Finkelstein, A., Funkhouser, T., Li, K., Liu, Z., . . . Wallace, G. (2001). Data Distribution Strategies for High-Resolution Displays. *Computers & Graphics,* 25(5), 811-818.
Cheng, C. I., & Wakefield, G. H. (1999). Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space*. Audio Engineering Society Convention 107*, New York.
Clark, H., Dutton, G., & Vanderlyn, P. (1957). The "Stereosonic" Recording and Reproducing System. *IRE Transactions on Audio,* 5(4), 96-111.
Codognet, P. (2003). Artificial Nature and Natural Artifice. In J. Shaw & P. Weibel (Eds.), *Future Cinema: The Cinematic Imaginary after Film* (pp. 462-465): MIT Press.
Cohen, A. J. (2001). Music as a Source of Emotion in Film. *Music and emotion: Theory and research*, 249-272.
Coleman, P. D. (1963). An Analysis of Cues to Auditory Depth Perception in Free Space. *Psychological Bulletin,* 60(3), 302.
Coleridge, S. T. (1817). *Biographia Literaria; or, Biographical Sketches of My Literary Life and Opinions*. New York: Published by Kirk and Merein.
Compton, P., Edwards, G., Kang, B., Lazarus, L., Malor, R., Menzies, T., . . . Sammut, C. (1991). Ripple Down Rules: Possibilities and Limitations*. Proceedings of the Sixth AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop, Calgary, Canada, University of Calgary*, 6-1.
Courchesne, L. (2005). *Panoramic and Horizontally Immersive Image Display System and Method*, U.S.A., Patent No. 6,905,218.
Courchesne, L. (2007). *Hemispherical Screen*, U.S.A., Patent No. D555,698.
Courchesne, L., Langlois, G., & Martinez, L. (2006). Where Are You?: An Immersive Experience in the Panoscope 360*. Proceedings of the 14th annual ACM international conference on Multimedia*, 1027-1028.

Couture, V., Langer, M., & Roy, S. (2010). Analysis of Disparity Distortions in Omnistereoscopic Displays. *ACM Transactions on Applied Perception (TAP),* 7(4), 25.

Couture, V., & Roy, S. (2013). The Omnipolar Camera: A New Approach to Stereo Immersive Capture. *2013 IEEE International Conference on Computational Photography (ICCP)*, Cambridge, MA.

Cowan, M. (2002). Digital Cinema Resolution–Current Situation and Future Requirements. http://etconsult.com/papers/Technical%20Issues%20in%20Cinema%20Resolution.pdf

Cremer, L. (1948). *Die Wissenschaftlichen Grundlagen Der Raumakustik*. Stuttgart,: S. Hirzel.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, 135-142.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communications of the ACM,* 35(6), 64-72.

Csikszentmihalyi, M. (1990). *Flow : The Psychology of Optimal Experience* (1st ed.). New York: Harper & Row.

Cummings, J. J., Bailenson, J. N., & Fidler, M. J. (2012, October 24-26). How Immersive Is Enough?: A Foundation for a Meta-Analysis of the Effect of Immersive Technology on Measured Presence. *International Society for Presence Research Annual Conference*, Philadelphia, Pennsylvania, USA.

Cutting, J. E. (1986a). *Perception with an Eye for Motion*: Mit Press Cambridge, MA.

Cutting, J. E. (1986b). The Shape and Psychophysics of Cinematic Space. *Behavior Research Methods, Instruments, & Computers,* 18(6), 551-558.

Cutting, J. E. (1987). Rigidity in Cinema Seen from the Front Row, Side Aisle. *Journal of Experimental Psychology: Human Perception and Performance,* 13(3), 323.

Cutting, J. E. (1997). How the Eye Measures Reality and Virtual Reality. *Behavior Research Methods, Instruments, & Computers,* 29(1), 27-36.

da Vinci, L. (1989). *Leonardo on Painting* (M. Kemp & M. Walker, Trans. M. Kemp Ed.): Yale University Press.

Daniel, J. (2003). Spatial Sound Encoding Including near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*.

Daniel, J., Nicol, R., & Moreau, S. (2003). Further Investigation of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging. *Audio Engineering Society Convention 114*, Amsterdam.

De La Gournerie, J. (1859). *Traité De Perspective Linéaire Contenant Les Tracés Pour Les Tableaux, Plans & Courbes, Les Bas Reliefs & Les Décorations Théatrales, Avec Une Théorie Des Effets De Perspective*: Dalmont et Dunod.

De Sena, E., Hacihabiboglu, H., & Cvetkovic, Z. (2013). Analysis and Design of Multichannel Systems for Perceptual Sound Field Reconstruction. *Audio, Speech, and Language Processing, IEEE Transactions on,* 21(8), 1653-1665.

DeFanti, T. A., Leigh, J., Renambot, L., Jeong, B., Verlo, A., Long, L., . . . Liu, Q. (2009). The Optiportal, a Scalable Visualization, Storage, and Computing Interface Device for the Optiputer. *Future Generation Computer Systems,* 25(2), 114-123.

Del Favero, D., & Barker, T. (2010). Scenario: Co-Evolution, Shared Autonomy and Mixed Reality. *Mixed and Augmented Reality-Arts, Media, and Humanities (ISMAR-AMH), 2010 IEEE International Symposium On*, 11-18.

Del Favero, D., Sewell, S., & Pagnucco, M. (2010). *Scenario* [AVIE-based artwork].

Dichgans, J., & Brandt, T. (1978). Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. *Perception* (pp. 755-804): Springer.

Dickerson, R. E., Kopka, M. L., & Pjura, P. (1983). A Random-Walk Model for Helix Bending in B-DNA. *Proceedings of the National Academy of Sciences,* 80(23), 7099-7103.

Dickins, G. (2003). *Soundfield Representation, Reconstruction and Perception.* (Master of Engineering Master), Australian National University. Retrieved from http://hdl.handle.net/1885/9728

Drummond, J. L. (1815). On Certain Appearances Observed in the Dissection of the Eyes of Fishes. *Transactions of the Royal Society of Edinburgh,* 12(7), 227-285.

Duguet, A.-M., Klotz, H., & Weibel, P. (1997). *Jeffrey Shaw-a User's Manual*. Karlsruhe: ZKM.

Ebert, D., Musgrave, F., Peachey, D., Perlin, K., & Worley, S. (2003). *Texturing & Modeling: A Procedural Approach* (3rd ed.): Morgan Kaufmann.

Ebert, R. (2010, May 9th 2010). Why I Hate 3D Movies. *Newsweek*.

Edgerton, S. Y. (1991). *The Heritage of Giotto's Geometry: Art and Science on the Eve of the Scientific Revolution*: Cornell University Press.

Eilemann, S., Makhinya, M., & Pajarola, R. (2009). Equalizer: A Scalable Parallel Rendering Framework. *Visualization and Computer Graphics, IEEE Transactions on,* 15(3), 436-452.

Einstein, A. (1905). On the Motion, Required by the Molecular-Kinetic Theory of Heat, of Particles Suspended in a Fluid at Rest. *Annalen der Physik,* 17, 149-560.

Enriquez, N. (2004). A Simple Construction of the Fractional Brownian Motion. *Stochastic processes and their applications,* 109(2), 203-223.

Falmier, O., & Young, M. E. (2008). The Impact of Object Animacy on the Appraisal of Causality. *The American Journal of Psychology*, 473-500.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What Causes the Face Inversion Effect? *J. Exp. Psychol. Hum. Percept. Perform.,* 21, 628-634.

Febretti, A., Nishimoto, A., Thigpen, T., Talandis, J., Long, L., Pirtle, J., . . . Plepys, D. (2013). CAVE2: A Hybrid Reality Environment for Immersive Simulation and Information Analysis. *IS&T/SPIE Electronic Imaging*, 864903-864903-864912.

Fischer, M., & Kornmüller, A. (1930). Optokinetisch Ausgelöste Bewegungswahrnehmung Und Optokinetischer Nystagmus [Optokinetically Induced Motion Perception and Optokinetic Nystagmus]. *Journal für Psychologie und Neurologie,* 41, 273-308.

Fisher, S. (1981). *Viewpoint Dependent Imaging: An Interactive Stereoscopic Display.* (Master of Science in Visual Studies Master of Science), MIT.

Fisher, S., Anderson, S., Ruiz, S., Naimark, M., Hoberman, P., Bolas, M., & Weinberg, R. (2005). Experiments in Interactive Panoramic Cinema. *Proc. SPIE 5664, Stereoscopic Displays and Virtual Reality Systems XII*, 626-632.

Flynn, B. (2012). *In Search of a Body: The Somatisation of Digital Cultural Heritage.* (PhD), University of New South Wales. Retrieved from http://handle.unsw.edu.au/1959.4/51614

Fontana, G. (1420). *Bellicorum Instrumentum Liber, Com Figuris Et Fictitys Litoris Conscriptus*. Bayerische Staatsbibliothek, Munich.

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & IJsselsteijn, W. (2000). Using Behavioral Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion Stimuli. *Presence: Teleoperators and Virtual Environments,* 9(2), 149-164.

Fröhlich, B., Blach, R., Stefani, O., Hochstrate, J., Hoffmann, J., Klüger, K., & Bues, M. (2005). Implementing Multi-Viewer Stereo Displays. *WSCG (Full Papers)*, 139-146.

Fulton, J. T. (2005). Processes in Biological Vision. http://www.4colorvision.com

Gaa, K., & Krüger, B. (1984). *Das Kaiser-Panorama. Berlin Um 1900. Bilder Aus Dem Berlin Der Jahrhundertwende.*: Berliner Festspiele.

Gardner, M. B. (1968). Proximity Image Effect in Sound Localization. *The Journal of the Acoustical Society of America,* 43, 163.

Gardner, W. G. (1997). Head Tracked 3-D Audio Using Loudspeakers. *Applications of Signal Processing to Audio and Acoustics, 1997. 1997 IEEE ASSP Workshop on*, 4 pp.

Gascón, G., Bayona, J. M., Espadero, J. M., & Otaduy, M. A. (2011). BlenderCAVE: Easy VR Authoring for Multi-Screen Displays. *SIACG 2011: Ibero-American Symposium in Computer Graphics*.

Gepshtein, S., & Banks, M. S. (2003). Viewing Geometry Determines How Vision and Haptics Combine in Size Perception. *Current Biology,* 13(6), 483-488.

Gerrig, R. J. (1993). *Experiencing Narrative Worlds : On the Psychological Activities of Reading*. New Haven: Yale University Press.

Gershun, A. (1939). The Light Field. *Journal of Mathematics and Physics,* XVIII, 51-151.

Gerz, J. (2000, 27 June). Jean-Michel Bruyère. *l'Humanité*.

Gibson, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.

Gibson, J. J. (1954). The Visual Perception of Objective Motion and Subjective Movement. *Psychological Review,* 61(5), 304.

Gibson, J. J. (1961). Ecological Optics. *Vision Research,* 1(3–4), 253-262.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1970). On Theories for Visual Space Perception. *Scandinavian Journal of Psychology,* 11(1), 75-79.

Gibson, J. J. (1971). The Information Available in Pictures. *Leonardo*, 27-35.

Gibson, J. J. (1972). A Theory of Direct Visual Perception. In J. Royce & W. Rozenboom (Eds.), *The Psychology of Knowing*. New York: Gordon & Breach.

Gibson, J. J. (1976). The Myth of Passive Perception: A Reply to Richards. *Philosophy and phenomenological research,* 37(2), 234-238.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilkey, R. H., Simpson, B. D., Isabelle, S. K., Kordik, A. J., & Weisenberger, J. M. (1999). Audition and the Sense of Presence in Virtual Environments. *The Journal of the Acoustical Society of America,* 105, 1163.

Gilkey, R. H., & Weisenberger, J. M. (1995). The Sense of Presence for the Suddenly Deafened Adult- Implications for Virtual Environments. *Presence: Teleoperators and Virtual Environments,* 4(4), 357-363.

Goldstein, E. B. (1981). The Ecology of J. J. Gibson's Perception. *Leonardo*, 191-195.

Gombrich, E. H. (1960). *Art and Illusion; a Study in the Psychology of Pictorial Representation*. New York: Pantheon Books.

Gortler, S. J., Grzeszczuk, R., Szeliski, R., & Cohen, M. F. (1996). The Lumigraph. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 43-54.

Grau, O. (1999). Into the Belly of the Image: Historical Aspects of Virtual Reality. *Leonardo,* 32(5), 365-371.

Grau, O. (2003). *Virtual Art: From Illusion to Immersion*: MIT press.

Grau, O. (2004). Immersion and Interaction. From Circular Frescoes to Interactive Image Spaces. *Survey of Media Art,* 1, 292-313.

Grau, O. (2007). *Mediaarthistories*: MIT Press (MA).

Grau, O., & Veigl, T. (2011). *Imagery in the 21st Century*: The MIT Press.

Greene, R. (1983). Determining the Preferred Viewpoint in Linear Perspective. *Leonardo*, 97-102.

Gregory, R. L. (1997). *Eye and Brain: The Psychology of Seeing*: Princeton university press.

Griesinger, D. (2002). Stereo and Surround Panning in Practice. *Audio Engineering Society Convention 112*, Munich.

Grimoin-Sanson, R. (1896a). *Appareil Appelé Le Phototachygraphe Sanson*, France, Patent No. 254,515.

Grimoin-Sanson, R. (1896b). *Projecteur Multiplex Dit Projecteur Sanson: Appareil Permettant De Pouvoir Prendre Et Projeter Des Vues Panoramiques Mouvementées Donnant L'illusion De La Réalité*, France, Patent No. 261,244.

Grimoin-Sanson, R. (1897). *Nouvel Appareil Permettant De Photographier Et De Projeter Sur Un Écran Circulaire Des Vues Animées Panoramiques En Couleur Par Le Cinécosmorama Sanson*, France, Patent No. 272,517.

Hammond, L. (1922). *Stereoscopic Motion Picture* U.S.A., Patent No. 1,435,520.

Harima, T., Shiga, K., Abe, K., Takane, S., Watanabe, K., & Sato, S. (2013). Relation between Perceived Direction of a Sound Image and the Behavior of the Precedence Effect. *Applied Acoustics,* 74(10), 1122-1135.

Hartigan, J. A. (1975). *Clustering Algorithms*: John Wiley & Sons, Inc.

Harvey, M. A., & Sanchez-Vives, M. V. (2005). The Binding Problem in Presence Research. *Presence: Teleoperators and Virtual Environments,* 14(5), 616-621.

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology,* 57(2), 243-259.

Heilig, M. (1960). *Stereoscopic-Television Apparatus for Individual Use*, U.S.A., Patent No. 2,955,156.

Held, R. T., & Banks, M. S. (2008). Misperceptions in Stereoscopic Displays: A Vision Science Perspective. *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, 23-32.

Hendrix, C., & Barfield, W. (1996). Presence within Virtual Environments as a Function of Visual Display Parameters. *Presence: Teleoperators and Virtual Environments,* 5(3), 274-289.

Hettinger, L. (2002). Illusory Self-Motion in Virtual Environments. *Handbook of virtual environments: Design, implementation, and applications*, 471-492.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–Accommodation Conflicts Hinder Visual Performance and Cause Visual Fatigue. *Journal of vision,* 8(3).

Hong, H., Jang, J., Lee, D., Lim, M., & Shin, H. (2010). Analysis of Angular Dependence of 3-D Technology Using Polarized Eyeglasses. *Journal of the Society for Information Display,* 18(1), 8-12.

Hopkins, A. A. (1898). *Magic: Stage Illusions and Scientific Diversions, Including Trick Photography*: Munn.

Horzu, P. (1954, March). Two Headed Tv Set. *Popular Science*.

Huang, H. C., & Hung, Y. P. (1998). Panoramic Stereo Imaging System with Automatic Disparity Warping and Seaming. *Graphical Models and Image Processing,* 60(3), 196-208.

Huang, H. C., Hung, Y. P., & Shih, S. W. (1996). Panoramic Stereo Imaging with Complete-Focus Views for Virtual Reality. *Proceeding of Second Workshop on Real-Time and Media Systems, RAMS'96*, p405-410.

Huang, K., Yuan, J.-C., Tsai, C.-H., Hsueh, W.-J., & Wang, N.-Y. (2003). How Crosstalk Affects Stereopsis in Stereoscopic Displays. *Electronic Imaging 2003*, 247-253.

Huhtamo, E. (2013). *Illusions in Motion: Media Archaeology of the Moving Panorama and Related Spectacles*: MIT Press.

Humphreys, G., Eldridge, M., Buck, I., Stoll, G., Everett, M., & Hanrahan, P. (2001). WireGL: A Scalable Graphics System for Clusters. *SIGGRAPH*, 1, 129-140.

Humphreys, G., Houston, M., Ng, R., Frank, R., Ahern, S., Kirchner, P. D., & Klosowski, J. T. (2002). Chromium: A Stream-Processing Framework for Interactive Rendering on Clusters. *ACM Transactions on Graphics (TOG)*, 21(3), 693-702.

Ijsselsteijn, W. A., Lombard, M., & Freeman, J. (2001). Toward a Core Bibliography of Presence. *Cyberpsychology & Behavior,* 4(2), 317-321.

Insu, Y., Mortensen, J., Khanna, P., Spanlang, B., & Slater, M. (2012). Visual Realism Enhances Realistic Response in an Immersive Virtual Environment - Part 2. *IEEE Computer Graphics and Applications,* 32(6), 36-45.

Ishiguro, H., Yamamoto, M., & Tsuji, S. (1990). Omni-Directional Stereo for Making Global Map. *Computer Vision, 1990. Proceedings, Third International Conference on*, 540-547.
Ishiguro, H., Yamamoto, M., & Tsuji, S. (1992). Omni-Directional Stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 14(2), 257-262.
ISPR. (2000). The Concept of Presence: Explication Statement. Accessed September 11 2013 at http://ispr.info/
Jack, C. E., & Thurlow, W. R. (1973). Effects of Degree of Visual Association and Angle of Displacement on the "Ventriloquism" Effect. *Perceptual and motor skills,* 37(3), 967-979.
Jacka, D., Reid, A., Merry, B., & Gain, J. (2007). A Comparison of Linear Skinning Techniques for Character Animation. *Proceedings of the 5th international conference on Computer graphics, virtual reality, visualisation and interaction in Africa*, 177-186.
Jacobs, G. H., Williams, G. A., Cahill, H., & Nathans, J. (2007). Emergence of Novel Color Vision in Mice Engineered to Express a Human Cone Photopigment. *Science,* 315(5819), 1723-1725.
Jacobson, J., Le Renard, M., Lugrin, J.-L., & Cavazza, M. (2005). The CaveUT System: Immersive Entertainment Based on a Game Engine. *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, 184-187.
Jacobson, J., & Lewis, M. (2005). Game Engine Virtual Reality with CaveUT. *Computer,* 38(4), 79-82.
James, D. L., & Twigg, C. D. (2005). Skinning Mesh Animations. *ACM Transactions on Graphics (TOG)*, 24(3), 399-407.
Jenkins, C. F. (1898). *Device for Obtaining Stereoscopic Effects in Exhibiting Pictures*, U.S.A., Patent No. 606,993.
Jin, C., Tan, T., Kan, A., Lin, D., von Schaik, A., Smith, K., & McGinity, M. (2005). Real-Time, Head-Tracked 3D Audio with Unlimited Simultaneous Sounds. *Proceedings of the 2005 International Conference on Auditory Display (ICAD)*, 1-4.
Johansson, G. (1973). Visual Perception of Biological Motion and a Model for Its Analysis. *Perception & psychophysics,* 14(2), 201-211.
Johnson, M. (1987). *The Body in the Mind : The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
Jorke, H., & Fritz, M. (2003). Infitec-a New Stereoscopic Visualisation Tool by Wavelength Multiplex Imaging. *Proceedings of Electronic Displays, September*.
Jorke, H., & Fritz, M. (2006). Stereo Projection Using Interference Filters. *Proc. SPIE 6055, Stereoscopic Displays and Virtual Reality Systems XIII*.
Jorke, H., & Simon, A. (2012). New High Efficiency Interference Filter Characteristics for Stereoscopic Imaging. *IS&T/SPIE Electronic Imaging*, 82881D-82881D-82889.
Jorke, H., Simon, A., & Fritz, M. (2008). Advanced Stereo Projection Using Interference Filters. *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, 2008*, 177-180.
Juarez, A., Schonenberg, W., & Bartneck, C. (2010). Implementing a Low-Cost CAVE System Using the Cryengine2. *Entertainment Computing,* 1(3), 157-164.
Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago University of Chicago Press
Kames, H. H. (1765). *Elements of Criticism* (3d ed.). n. p.
Kan, A., Jin, C., Tan, T., Lin, D., van Schaik, A., Smith, K., & McGinity, M. (2005). 3dape: A Real-Time 3D Audio Playback Engine. *Proceedings of AES 118th Convention. Barcelona:[sn]*, 6, 343.
Kendall, G. S. (1995). The Decorrelation of Audio Signals and Its Impact on Spatial Imagery. *Computer Music Journal,* 19(4), 71-87.

Kenderdine, S. (2010). Immersive Visualization Architectures and Situated Embodiments of Culture and Heritage. *Information Visualisation (IV), 2010 14th International Conference*, 408-414.
Kenderdine, S., Forte, M., & Camporesi, C. (2011). Rhizome of Western Han: An Omnispatial Theatre for Archaeology. *Revive the Past*, 141.
Kenderdine, S., Shaw, J., Del Favero, D., & Brown, N. (2007). Place-Hampi: Co-Evolutionary Narrative & Augmented Stereographic Panoramas. *New Heritage: New Media And Cultural Heritage, Vijayanagara, India, Routledge, London*.
Kim, S.-H., Feldman, J., & Singh, M. (2013). Perceived Causality Can Alter the Perceived Trajectory of Apparent Motion. *Psychological science,* 24(4), 575-582.
Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial Updating of Self-Position and Orientation During Real, Imagined, and Virtual Locomotion. *Psychological science,* 9(4), 293-298.
Kober, S. E., Kurzmann, J., & Neuper, C. (2012). Cortical Correlate of Spatial Presence in 2d and 3D Interactive Virtual Reality: An EEG Study. *International Journal of Psychophysiology,* 83(3), 365-374.
Können, G. (1985). *Polarized Light in Nature*: CUP Archive.
Kooi, F. L., & Toet, A. (2004). Visual Comfort of Binocular and 3D Displays. *Displays,* 25(2), 99-108.
Kottenhoff, H. (1961). *Was Ist Richtiges Sehen Mit Umkehrbrillen Und in Welchem Sinne Stellt Sich Das Sehen Um?* (Vol. 5): A. Hain.
Krueger, M. W. (1977, June 13-16). Responsive Environments*. Proceedings of the 1977 National Computer Conference*, Dallas, Texas, 423-433.
Krueger, M. W., Gionfriddo, T., & Hinrichsen, K. (1985). VIDEOPLACE - an Artificial Reality. *SIGCHI Bull.,* 16(4), 35-40.
Kubitz, W., & Poppelbaum, W. (1973). Stereomatrix, an Interactive Three Dimensional Computer Display*. Society for Information Display, Proceedings*, 14, 94-98.
Kubovy, M. (1988). *The Psychology of Perspective and Renaissance Art*: CUP Archive.
Kuchelmeister, V., Shaw, J., McGinity, M., Del Favero, D., & Hardjono, A. (2009). Immersive Mixed Media Augmented Reality Applications and Technology. *Advances in Multimedia Information Processing-PCM 2009* (pp. 1112-1118): Springer.
Kuhlen, T., Assenmacher, I., & Lentz, T. (2007). A True Spatial Sound System for CAVE-Like Displays Using Four Loudspeakers. *Virtual Reality* (pp. 270-279): Springer.
Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., & Froehlich, B. (2011). C1x6: A Stereoscopic Six-User Display for Co-Located Collaboration in Shared Virtual Environments. *ACM Transactions on Graphics,* 30(6), 1-12.
Lake, A. (1893). *Illusion Apparatus*, U.S.A., Patent No. 508,227.
Lambert, J.-H. (1760). *Photometria, Sive De Mensura Et Gradibus Luminis, Colorum Et Umbrae*. Augsburg: sumptibus viduae E. Klett.
Land, E. H. (1937). *Polarizing Optical System*, U.S.A., Patent No. 2,099,694.
Land, E. H. (1940). *Apparatus Employing Polarized Light for the Production of Stereoscopic Images*, U.S.A., Patent No. 2,203,687.
Land, E. H. (1942a). *Composite Film for Receiving Stereo*, U.S.A., Patent No. 2,289,715.
Land, E. H. (1942b). *Light-Polarizing Image in Full Color*, U.S.A., Patent No. 2,289,714.
Laurel, B. (1991). *Computers as Theatre*. Reading, Mass.: Addison-Wesley Pub.
Lee, K. M. (2004). Presence, Explicated. *Communication Theory,* 14(1), 27-50.
Leguey, C. F. A., & Bap, F. P. G. (1902a). *Appareil Pour La Production Des Films Destinés Aux Panoramas Cinématographiques*, France, Patent No. 327,058.
Leguey, C. F. A., & Bap, F. P. G. (1902b). *Panorama Cinématographique*, France, Patent No. 327,090.

Lepecq, J.-C., Bringoux, L., Pergandi, J.-M., Coyle, T., & Mestre, D. (2009). Afforded Actions as a Behavioral Assessment of Physical Presence in Virtual Environments. *Virtual reality,* 13(3), 141-151.

Levoy, M., & Hanrahan, P. (1996). Light Field Rendering. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 31-42.

Lewis, J. P., Cordner, M., & Fong, N. (2000). Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-Driven Deformation. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 165-172.

Linden, D. E., Kallenbach, U., Heinecke, A., Singer, W., & Goebel, R. (1999). The Myth of Upright Vision. A Psychophysical and Functional Imaging Study of Adaptation to Inverting Spectacles. *PERCEPTION-LONDON-,* 28, 469-482.

Lippmann, G. (1908). La Photographie Intergale. *Comptes Rendus de l'Académie des Sciences,* 146(9), 446-451.

Lipton, L. (2012). Brief History of Electronic Stereoscopic Displays. *Optical Engineering,* 51(2), 021103-021101-021103-021105.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The Precedence Effect. *The Journal of the Acoustical Society of America,* 106(4), 22.

Litovsky, R. Y., & Shinn-Cunningham, B. G. (2001). Investigation of the Relationship among Three Common Measures of Precedence: Fusion, Localization Dominance, and Discrimination Suppression. *Journal of the Acoustical Society of America,* 109(1), 346-358.

Liu, Y.-W., & Smith III, J. O. (2002). Perceptually Similar Orthogonal Sounds and Applications to Multichannel Acoustic Echo Canceling. *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication,* 3(2), 0-0.

Lombard, M., & Jones, M. T. (2006). Defining Presence. *Ninth International Workshop on Presence*, Cleveland, Ohio.

Lombard, M., & Jones, M. T. (2007). Identifying the (Tele) Presence Literature. *PsychNology Journal,* 5(2), 197-206.

Lossius, T. (2007). *Sound Space Body: Reflections on Artistic Practice* Retrieved from http://www.khib.no/media/1566962/KR_Trond_Lossius.pdf

Lucretius, & Leonard, W. E. (1921). *Of the Nature of Things*: JM Dent & Sons.

Lugrin, J.-L., Charles, F., Cavazza, M., Le Renard, M., Freeman, J., & Lessiter, J. (2012). CaveUDK: A VR Game Engine Middleware. *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, 137-144.

Mach, E. (1875). *Grundlinien Der Lehre Von Den Bewegungsempfindungen*: Engelmann.

Majumder, A., & Stevens, R. (2004). Color Nonuniformity in Projection-Based Displays: Analysis and Solutions. *IEEE Transactions on Visualization and Computer Graphics,* 10(2), 177-188.

Malus, E.-L. (1809). Sur Une Propriété De La Lumière Réfléchie. *Memoires de Physique et de Chimie de la Societe d 'Arcueil,* 2, 143-158.

Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. New York: WH Freeman and Co.

Mania, K., & Robinson, A. (2004). The Effect of Quality of Rendering on User Lighting Impressions and Presence in Virtual Environments. *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry*, 200-205.

Manilius. (1977). *Astronomica* (G. P. Goold, Trans.). Cambridge: Harvard University Press.

Mather, G., & West, S. (1993). Recognition of Animal Locomotion from Dynamic Point-Light Displays. *Perception,* 22, 759-766.

McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural Substrates of Perceptual Enhancement by Cross-Modal Spatial Attention. *Journal of Cognitive Neuroscience,* 15(1), 10-19.

Plausible Blinn-Phong Reflection of Standard Cube MIP-Maps (2013). Plausible Blinn-Phong Reflection of Standard Cube MIP-Maps (pp. 8): Williams College Department of Computer Science.

McMillan, L., & Bishop, G. (1995). Plenoptic Modeling: An Image-Based Rendering System. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, 39-46.

McQuire, S., & Papastergiadis, N. (2006). *Conversations: The Parallax Effect*: University of New South Wales, College of Fine Arts.

Melvill Jones, G., Guitton, D., & Berthoz, A. (1988). Changing Patterns of Eye-Head Coordination During 6 Hours of Optically Reversed Vision. *Experimental Brain Research,* 69(3), 531-544.

Mercier, A. (1998). Portefeuille Industriel: Auguste Baron Et Le Cinématographe - II. Le Cinématorama Et Le Cinépanorama *La Revue, Musée des Arts et Metiers,* 24.

Mershon, D. H., Desaulniers, D. H., Amerson, T. L., & Kiefer, S. A. (1980). Visual Capture in Auditory Distance Perception: Proximity Image Effect Reconsidered. *Journal of Auditory Research*.

Michotte, A. (1963). *The Perception of Causality*. Oxford: Basic Books.

Minsky, M. (1980). Telepresence. *Omni,* 2(9), 45-51.

Moldrzyk, C., Goertz, A., Makarski, M., Feistel, S., Ahnert, W., & Weinzierl, S. (2007). Wellenfeldsynthese Fur Einen Grossen Horsaal. *FORTSCHRITTE DER AKUSTIK,* 33(2), 683.

Montrym, J. S., Baum, D. R., Dignam, D. L., & Migdal, C. J. (1997). InfiniteReality: A Real-Time Graphics System. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, 293-302.

Mörters, P., & Peres, Y. (2010). *Brownian Motion* (Vol. 30): Cambridge University Press.

Mossio, M., & Taraborelli, D. (2008). Action-Dependent Perceptual Invariants: From Ecological to Sensorimotor Approaches. *Consciousness and cognition,* 17(4), 1324-1340.

Murray, C. D., Arnold, P., & Thornton, B. (2000). Presence Accompanying Induced Hearing Loss: Implications for Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments,* 9(2), 137-148.

Musgrave, F. K., Kolb, C. E., & Mace, R. S. (1989). The Synthesis and Rendering of Eroded Fractal Terrains. *ACM SIGGRAPH Computer Graphics*, 23(3), 41-50.

Naemura, T., Kaneko, M., & Harashima, H. (1998). Multi-User Immersive Stereo. *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, 1, 903-907.

Naemura, T., Takano, T., Kaneko, M., & Harashima, H. (1997). Ray-Based Creation of Photo-Realistic Virtual World. *Virtual Systems and MultiMedia, 1997. VSMM'97. Proceedings., International Conference on*, 59-68.

Naimark, M. (2005). Two Unusual Projection Spaces. *Presence: Teleoperators and Virtual Environments,* 14(5), 597-506.

Naimark, M., & Felsenstein, L. (1997). *Panoramic Display with Stationary Display Device and Rotating Support Structure*, U.S.A., Patent No. US5601353A.

Navrátil, P. A., Westing, B., Johnson, G. P., Athalye, A., Carreno, J., & Rojas, F. (2009). A Practical Guide to Large Tiled Displays. *Advances in Visual Computing* (pp. 970-981). Berlin: Springer.

Nawrot, M., & Stroyan, K. (2009). The Motion/Pursuit Law for Visual Depth Perception from Motion Parallax. *Vision Research,* 49(15), 1969-1978.

Transmission of Presence (1980). Transmission of Presence *MIT Architecture Machine Group proposal to Cybernetics Technology Dividion of Defence Advanced Research Projects Agency*. Cambridge: MIT.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology*: WH Freeman/Times Books/Henry Holt & Co.

Neumann, J., Fermüller, C., & Aloimonos, Y. (2004). A Hierarchy of Cameras for 3D Photography. *Computer Vision and Image Understanding,* 96(3), 274-293.

Nicodemus, F. E. (1965). Directional Reflectance and Emissivity of an Opaque Surface. *Applied Optics,* 4(7), 767-773.

Njami, S., Maurin, S., & Kouoh, K. (2013). *Issa Samb: Word! Word? Word! And the Undecipherable Form* (K. Kouoh Ed.): Sternberg Press.

Noë, A. (2004). *Action in Perception*: MIT Press.

Ono, H. (2008, 15-16 Dec. 2008). Perception of Depth, Motion, and Stability with Motion Parallax (Invited Paper). *Universal Communication, 2008. ISUC '08. Second International Symposium on*, 193-198.

Ordine, N. (1996). *Giordano Bruno and the Philosophy of the Ass*: Yale University Press.

Oren, M., & Nayar, S. K. (1994). Generalization of Lambert's Reflectance Model. *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, 239-246.

Ovid, & Humphries, R. (1955). *Metamorphoses* (R. Humphries, Trans. Vol. 1983). Bloomington: Indiana University Press.

Ovid, & Innes, M. M. (1955). *The Metamorphoses of Ovid* (M. M. Innes, Trans.): Penguin Books.

Ovid, & Kline, A. S. (2000). *Metamorphoses* (A. S. Kline, Trans.).

Ovid, & Sandys, G. (1632). *Metamorphoses* (G. Sandys, Trans.). Oxford: John Lichfield.

Pajdla, T. (2002). Stereo with Oblique Cameras. *International Journal of Computer Vision,* 47(1-3), 161-170.

Pala, S., Stevens, R., & Surman, P. (2007). Optical Crosstalk and Visual Comfort of a Stereoscopic Display Used in a Real-Time Application. *Proc. SPIE 6490, Stereoscopic Displays and Virtual Reality Systems XIV*.

Scientific American Supplement (1900). The Panoramas of the Paris Exposition (Vol. 1287, pp. 2).

Papadopoulos, C., Petkov, K., & Kaufman, A. E. (2013). Building the Reality Deck. *ACM SIGCHI Conference on Human Factors in Computing Systems*, Paris.

Park, H., Lee, M.-H., Kim, S.-J., & Park, J.-I. (2005). Specularity-Free Projection on Nonplanar Surface. *Advances in Multimedia Information Processing-PCM 2005* (pp. 606-616): Springer.

Pastoor, S. (1995). Human Factors of 3D Imaging: Results of Recent Research at Heinrich-Hertz-Institut Berlin. *Proc. IDW*, 95(3), 69-72.

Pearson, K. (1905). The Problem of the Random Walk. *Nature,* 72(1865), 294.

Peleg, S., & Ben-Ezra, M. (1999). Stereo Panorama with a Single Camera. *IEEE Computer Vision and Pattern Recognition, 1999. Proceedings.*, 1.

Peleg, S., Ben-Ezra, M., & Pritch, Y. (2000). Cameras for Stereo Panoramic Imaging. *IEEE Computer Vision and Pattern Recognition, 2000. Proceedings.*, 1, 208-214.

Peleg, S., Ben-Ezra, M., & Pritch, Y. (2001). Omnistereo: Panoramic Stereo Imaging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 23(3), 279-290.

Peleg, S., & Herman, J. (1997). Panoramic Mosaics by Manifold Projection. *IEEE Computer Vision and Pattern Recognition, 1997. Proceedings.*, 338-343.

Perkins, D. N. (1973). Compensating for Distortion in Viewing Pictures Obliquely. *Perception & Psychophysics,* 14(1), 13-18.

Perlin, K. (1985). An Image Synthesizer. *ACM SIGGRAPH Computer Graphics,* 19(3), 287-296.

Perlin, K. (2002). Improving Noise. *ACM Transactions on Graphics (TOG)*, 21(3), 681-682.

Perlin, K., & Hoffert, E. M. (1989). Hypertexture. *ACM SIGGRAPH Computer Graphics*, 23(3), 253-262.
Philbeck, J., & Mershon, D. (2002). Knowledge About Typical Source Output Influences Perceived Auditory Distance. *The Journal of the Acoustical Society of America,* 111(5 Pt 1), 1980-1983.
Phong, B. T. (1975). Illumination for Computer Generated Pictures. *Communications of the ACM,* 18(6), 311-317.
Scientific American (1902). The Photorama (Vol. 86, pp. 344-345).
Pirenne, M. H. (1970). *Optics, Painting & Photography*: Cambridge UP.
Poletti, M. A. (2000). A Unified Theory of Horizontal Holographic Sound Systems. *Journal of the Audio Engineering Society,* 48(12), 1155-1182.
Pollock, B., Burton, M., Kelly, J. W., Gilbert, S., & Winer, E. (2012). The Right View from the Wrong Location: Depth Perception in Stereoscopic Multi-User Virtual Environments. *Visualization and Computer Graphics, IEEE Transactions on,* 18(4), 581-588.
Potard, G., & Burnett, I. (2004). Decorrelation Techniques for the Rendering of Apparent Sound Source Width in 3d Audio Displays. *Proc. of the 7th Int. Conf. on Digital Audio Effects*, 5-8.
Powers, W. T. (1973). *Behavior: The Control of Perception*. New York: Hawthorne.
Premecz, M. (2006). Iterative Parallax Mapping with Slope Information. *Central European Seminar on Computer Graphics*.
Pritch, Y., Ben-Ezra, M., & Peleg, S. (2000). Automatic Disparity Control in Stereo Panoramas (Omnistereo). *Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on*, 54-61.
Pross, A., Blach, R., Bues, M., Reichel, R., & Stefani, O. (2012, February 9th). Optimization of a Multi-View System Based on Pulsed Led-LCD Projectors. *Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII*.
Prothero, J. D. (1998). *The Role of Rest Frames in Vection, Presence and Motion Sickness.* (PhD), University of Washington.
Prothero, J. D., Hoffman, H. G., Parker, D. E., Furness, T. A., & Wells, M. J. (1995). Foreground/Background Manipulations Affect Presence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 39(21), 1410-1414.
Ptito, M., Faubert, J., Gjedde, A., & Kupers, R. (2003). Separate Neural Pathways for Contour and Biological-Motion Cues in Motion-Defined Animal Shapes. *Neuroimage,* 19(2), 246-252.
Pulkki, V. (1997). Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *Journal of the Audio Engineering Society,* 45(6), 456-466.
Pulkki, V. (1999). Uniform Spreading of Amplitude Panned Virtual Sources. *Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop on*, 187-190.
Pulkki, V. (2001a). Localization of Amplitude-Panned Virtual Sources II: Two-and Three-Dimensional Panning. *Journal of the Audio Engineering Society,* 49(9), 753-767.
Pulkki, V. (2001b). *Spatial Sound Generation and Perception by Amplitude Panning Techniques.* (PhD), Helsinki University of Technology.
Pulkki, V. (2002). Compensating Displacement of Amplitude-Panned Virtual Sources. *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*.
Pulkki, V., & Karjalainen, M. (2001). Localization of Amplitude-Panned Virtual Sources I: Stereophonic Panning. *Journal of the Audio Engineering Society,* 49(9), 739-752.
Pulkki, V., & Karjalainen, M. (2008). Multichannel Audio Rendering Using Amplitude Panning [Dsp Applications]. *Signal Processing Magazine, IEEE,* 25(3), 118-122.
Pulkki, V., Karjalainen, M., & Huopaniemi, J. (1999). Analyzing Virtual Sound Source Attributes Using a Binaural Auditory Model. *Journal of the Audio Engineering Society,* 47(4), 203-217.

Purkinje, J. E. (1825). *Neue Beiträge Zur Kenntniss Des Sehens in Subjektiver Hinsicht*. Berlin: Reimer.

Rabenstein, R., & Spors, S. (2005). Spatial Sound Reproduction with Wave Field Synthesis. *Atti del Congresso*.

Ramos-Fernandez, G., Mateos, J. L., Miramontes, O., Cocho, G., Larralde, H., & Ayala-Orozco, B. (2004). Lévy Walk Patterns in the Foraging Movements of Spider Monkeys (Ateles Geoffroyi). *Behavioral Ecology and Sociobiology,* 55(3), 223-230.

Read, J. (2006). Checking the “Ghosting” between Projectors. Accessed August 10 2013 at http://www.jennyreadresearch.com/research/lab-set-up/checking-the-ghosting-between-the-projectors/

Reed, E. S. (1988). *James J. Gibson and the Psychology of Perception*: Yale University Press New Haven, CT.

Reuschel, J., Drewing, K., Henriques, D. Y., Rösler, F., & Fiehler, K. (2010). Optimal Integration of Visual and Proprioceptive Movement Information for the Perception of Trajectory Geometry. *Experimental brain research,* 201(4), 853-862.

Revonsuo, A. (1999). Binding and the Phenomenal Unity of Consciousness. *Consciousness and cognition,* 8(2), 173-185.

Rey, B., Alcañiz, M., Tembl, J., & Parkhutik, V. (2010). Brain Activity and Presence: A Preliminary Study in Different Immersive Conditions Using Transcranial Doppler Monitoring. *Virtual Reality,* 14(1), 55-65.

Rhee, I., Shin, M., Hong, S., Lee, K., Kim, S. J., & Chong, S. (2011). On the Levy-Walk Nature of Human Mobility. *IEEE/ACM Transactions on Networking (TON),* 19(3), 630-643.

Richards, M., & Gomes, G. (2011). *Spectral Separation Filters for 3D Stereoscopic D-Cinema Presentation*, U.S.A., Patent No. 7,959,295 B2.

Riecke, B. (2006). Simple User-Generated Motion Cueing Can Enhance Self-Motion Perception (Vection) in Virtual Reality. *Proceedings of the ACM symposium on Virtual reality software and technology*, 104-107.

Riecke, B. (2010). Compelling Self-Motion through Virtual Environments without Actual Self-Motion: Using Self-Motion Illusions (“Vection”) to Improve User Experience in VR. *Virtual reality*, 149-176.

Riecke, B., Cunningham, D., & Bülthoff, H. (2007). Spatial Updating in Virtual Reality: The Sufficiency of Visual Information. *Psychological research,* 71(3), 298-313.

Riecke, B., Feuereissen, D., Rieser, J., & McNamara, T. (2011). Spatialized Sound Enhances Biomechanically-Induced Self-Motion Illusion (Vection). *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2799-2802.

Riecke, B., Feuereissen, D., Rieser, J., & McNamara, T. (2012, 4-8 March 2012). Self-Motion Illusions (Vection) in VR - Are They Good for Anything? *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, 35-38.

Riecke, B., Heyde, M. V. D., & Bülthoff, H. (2005). Visual Cues Can Be Sufficient for Triggering Automatic, Reflexlike Spatial Updating. *ACM Transactions on Applied Perception (TAP),* 2(3), 183-215.

Riecke, B., & Schulte-Pelkum, J. (2013). Perceptual and Cognitive Factors for Self-Motion Simulation in Virtual Environments: How Can Self-Motion Illusions (“Vection”) Be Utilized? *Human Walking in Virtual Environments* (pp. 27-54): Springer.

Riecke, B., Schulte-Pelkum, J., Avraamides, M., & Bülthoff, H. (2004). Enhancing the Visually Induced Self-Motion Illusion (Vection) under Natural Viewing Conditions in Virtual Reality. *Proceedings of Seventh Annual Workshop Presence 2004*, 125-132.

Riecke, B., Schulte-Pelkum, J., & Caniard, F. (2006). Using the Perceptually Oriented Approach to Optimize Spatial Presence & Ego-Motion Simulation. *Handbook of Presence. Hillsdale: Lawrence Erlbaum (submitted)*.

Riecke, B., Sigurdarson, S., & Milne, A. (2012). Moving through Virtual Reality without Moving? *Cognitive Processing,* 13(1), 293-297.
Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., . . . Alcañiz, M. (2007). Affective Interactions Using Virtual Reality: The Link between Presence and Emotions. *CyberPsychology & Behavior,* 10(1), 45-56.
Rogers, B., & Graham, M. (1979). Motion Parallax as an Independent Cue for Depth Perception. *Perception,* 8(2), 125-134.
Rogers, B., & Graham, M. (1982). Similarities between Motion Parallax and Stereopsis in Human Depth Perception. *Vision research,* 22(2), 261-270.
Rokeby, D. (1995). Transforming Mirrors. *Leonardo Electronic Almanac,* 3(4), 12.
Rokeby, D. (1998). The Construction of Experience: Interface as Content. *Digital Illusion: Entertaining the future with high technology*, 27-48.
Rollmann, W. (1853). Zwei Neue Stereoskopische Methoden. *Annalen der Physik,* 166(9), 186-187.
Rosinski, R. R., & Farber, J. (1980). Compensation for Viewing Point in the Perception of Pictured Space. In M. A. Hagen (Ed.), *The Perception of Pictures*. New York: Academic.
Rouch, J., & Feld, S. (2003). *Ciné Ethnography*: University of Minnesota Press.
Ryan, M.-L. (2001). *Narrative as Virtual Reality*: Johns Hopkins University Press Baltimore, MA.
Salomon, D. (2006). *Transformations and Projections in Computer Graphics*: Springer.
Sambo, C. F., & Forster, B. (2009). An ERP Investigation on Visuotactile Interactions in Peripersonal and Extrapersonal Space: Evidence for the Spatial Rule. *Journal of cognitive neuroscience,* 21(8), 1550-1559.
Sanchez-Vives, M. V., & Slater, M. (2005). From Presence to Consciousness through Virtual Reality. *Nature Reviews Neuroscience,* 6(4), 332-339.
Sarachik, K. B. (1989). Characterising an Indoor Environment with a Mobile Robot and Uncalibrated Stereo*. Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*, 984-989.
Scafetta, N. (2011). Understanding the Complexity of the Lévy-Walk Nature of Human Mobility with a Multi-Scale Cost/Benefit Model. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 21(4), 043106-043106-043110.
Schaeffer, B., & Goudeseune, C. (2003). Syzygy: Native PC Cluster VR*. Proceedings of IEEE Virtual Reality 2003*, 15-22.
Schafer, D. (1982, June 29). Stereo Painting Display Devices*. Visual Simulation and Image Realism II*, 78.
Schloerb, D. W. (1995). A Quantitative Measure of Telepresence. *Presence: Teleoperators and Virtual Environments,* 4(1), 64-80.
Scholl, B. J., & Nakayama, K. (2002). Causal Capture: Contextual Effects on the Perception of Collision Events. *Psychological Science,* 13(6), 493-498.
Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual Causality and Animacy. *Trends in cognitive sciences,* 4(8), 299-309.
Schubert, T. W. (2009). A New Conception of Spatial Presence: Once Again, with Feeling. *Communication Theory,* 19(2), 161-187.
Schulze, J. P., Prudhomme, A., Weber, P., & DeFanti, T. A. (2013). CalVR: An Advanced Open Source Virtual Reality Software Framework*. IS&T/SPIE Electronic Imaging*, 864902-864902-864908.
Seitz, S. M., & Kim, J. (2002). The Space of All Stereo Images. *International Journal of Computer Vision,* 48(1), 21-38.
Seuntiëns, P., Meesters, L., & IJsselsteijn, W. (2005). Perceptual Attributes of Crosstalk in 3D Images. *Displays,* 26(4), 177-183.

Sharp, G. D., Coleman, D., Gilman, S., & Chen, J. (2013). Achieving High Stereo-Contrast-Ratio in Polarization-Based 3D Front Projection. *SID Symposium Digest of Technical Papers*, 44(1), 744-747.

Shaw, J., & Catala, L. (2013). Jeffrey Shaw, the Art of Interactive Cinema. *Digitalarti, 13*.

Shaw, J., & Kenderdine, S. (2006). *Place-Hampi* [AVIE-based artwork].

Sherman, W. R., Coming, D., & Su, S. (2013). FreeVR: Honoring the Past, Looking to the Future. *IS&T/SPIE Electronic Imaging*, 864906-864906-864915.

Shestak, S., Kim, D., & Kim, Y. (2012). How Much Crosstalk Can Be Allowed in a Stereoscopic System at Various Grey Levels? *Stereoscopic Displays and Applications XXIII, Proc. SPIE,* 8288, 828810.

Shibata, T., Kim, J., Hoffman, D. M., & Banks, M. S. (2011). The Zone of Comfort: Predicting Visual Discomfort with Stereo Displays. *Journal of vision,* 11(8).

Shimamura, J., Takemura, H., Yokoya, N., & Yamazawa, K. (2000a). Construction and Presentation of a Virtual Environment Using Panoramic Stereo Images of a Real Scene and Computer Graphics Models. *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 4, 463-467.

Shimamura, J., Takemura, H., Yokoya, N., & Yamazawa, K. (2000b). Construction of an Immersive Mixed Environment Using an Omnidirectional Stereo Image Sensor. *Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on*, 62-69.

Shum, H.-Y., & He, L.-W. (1999). Rendering with Concentric Mosaics. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 299-306.

Shum, H.-Y., Kalai, A., & Seitz, S. M. (1999). Omnivergent Stereo. *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1, 22-29.

Shum, H.-Y., Ng, K.-T., & Chan, S.-C. (2000). Virtual Reality Using the Concentric Mosaic: Construction, Rendering and Data Compression. *Image Processing, 2000. Proceedings. 2000 International Conference on*, 3, 644-647.

Simion, F., Regolin, L., & Bulf, H. (2008). A Predisposition for Biological Motion in the Newborn Baby. *Proceedings of the National Academy of Sciences,* 105(2), 809-813.

Simon, A., & Göbel, M. (2002). The I-Cone™ – a Panoramic Display System for Virtual Environments. *Pacific Graphics '02*.

Simon, A., & Jorke, H. (2011). Interference Filter System for High-Brightness and Natural-Color Stereoscopic Imaging. *SID Symposium Digest of Technical Papers,* 42(1), 317-319.

Simon, A., Smith, R. C., & Pawlicki, R. R. (2004). Omnistereo for Panoramic Virtual Environment Display Systems. *Proceedings of IEEE Virtual Reality 2004*, 67-279.

Simpson, B. D., Hale, D. W., Isabelle, S. K., & Gilkey, R. H. (1996). The Experiences of Untrained Subjects Listening to Virtual Sounds. *The Journal of the Acoustical Society of America,* 100, 2633.

Sims, D. W., Humphries, N. E., Bradford, R. W., & Bruce, B. D. (2012). Lévy Flight and Brownian Search Patterns of a Free-Ranging Predator Reflect Different Prey Field Characteristics. *Journal of Animal Ecology,* 81(2), 432-442.

Slater, M. (2003). A Note on Presence Terminology. *Presence connect,* 3(3).

Slater, M. (2009). Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philos Trans R Soc Lond B Biol Sci,* 364(1535), 3549-3557.

Slater, M., Khanna, P., Mortensen, J., & Yu, I. (2009). Visual Realism Enhances Realistic Response in an Immersive Virtual Environment. *IEEE Computer Graphics and Applications,* 29(3), 76-84.

Slater, M., McCarthy, J., & Maringelli, F. (1998). The Influence of Body Movement on Subjective Presence in Virtual Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 40(3), 469-477.

Slater, M., Spanlang, B., & Corominas, D. (2010). Simulating Virtual Environments within Virtual Environments as the Basis for a Psychophysics of Presence. *ACM Transactions on Graphics (SIGGRAPH),* 29(4), 1-9.

Slater, M., Usoh, M., & Chrysanthou, Y. (1995). The Influence of Dynamic Shadows on Presence in Immersive Virtual Environments. *Virtual Environments' 95* (pp. 8-21): Springer.

SMPTE 431-1-2006 D-Cinema Quality - Screen Luminance Level, Chromaticity and Uniformity, SMPTE, 2006

Soares, L., Raffin, B., & Jorge, J. (2008). PC Clusters for Virtual Reality. *IJVR,* 7(1), 67-80.

Soares, L. P., Raposo, A., Jorge, J., Araujo, B., Dias, M., & Carvalho, F. (2010, Aug. 30 2010-Sept. 3 2010). Multi-Projector VR Systems*. Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2010 23rd SIBGRAPI Conference on*, 10-21.

Spector, R. H. (1990). Visual Fields. *Clinical Methods: The History, Physical, and Laboratory Examinations* (3rd ed., pp. 565-572). Boston: Butterworths.

Spors, S., & Rabenstein, R. (2006). Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Arrays Used for Wave Field Synthesis *Audio Engineering Society Convention 120*, Paris.

Spors, S., Rabenstein, R., & Ahrens, J. (2008). The Theory of Wave Field Synthesis Revisited*. Audio Engineering Society Convention 124*, Amsterdam, 17-20.

Spors, S., Teutsch, H., Kuntz, A., & Rabenstein, R. (2004). Sound Field Synthesis. In Y. Huang & J. Benesty (Eds.), *Audio Signal Processing for Next-Generation Multimedia Communication Systems* (pp. 344): Springer.

Sridhar, A. (2012). *Combining Computer Vision and Knowledge Acquisition to Provide Real-Time Activity Recognition for Multiple Persons within Immersive Environments.* (PhD), University of New South Wales, Sydney.

Sridhar, A., & Sowmya, A. (2008). Multiple Camera, Multiple Person Tracking with Pointing Gesture Recognition in Immersive Environments. *Advances in Visual Computing* (pp. 508-519): Springer.

Sridhar, A., & Sowmya, A. (2009). SparseSPOT: Using a Priori 3-D Tracking for Real-Time Multi-Person Voxel Reconstruction*. Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, 135-138.

Sridhar, A., & Sowmya, A. (2011). Distributed, Multi-Sensor Tracking of Multiple Participants within Immersive Environments Using a 2-Cohort Camera Setup. *Machine Vision and Applications,* 22(6), 1009-1025.

Sridhar, A., Sowmya, A., & Compton, P. (2010). On-Line, Incremental Learning for Real-Time Vision Based Movement Recognition*. Machine Learning and Applications (ICMLA), 2010 Ninth International Conference on*, 465-470.

Stockman, A., & Sharpe, L. T. (2006). Into the Twilight Zone: The Complexities of Mesopic Vision and Luminous Efficiency. *Ophthalmic and Physiological Optics,* 26(3), 225-239.

Stoll, G., Eldridge, M., Patterson, D., Webb, A., Berman, S., Levy, R., . . . Hanrahan, P. (2001). Lightning-2: A High-Performance Display Subsystem for PC Clusters*. Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 141-148.

Storms, R. L. (2002). Auditory-Visual Cross-Modality Interaction and Illusions. *Handbook of virtual environments: Design, implementation, and applications*, 455-470.

Stratton, G. M. (1897). Vision without Inversion of the Retinal Image. *Psychological review,* 4(4), 341.

Streit, A., Christie, R., & Boud, A. (2004). Understanding Next-Generation VR: Classifying Commodity Clusters for Immersive Virtual Reality*. Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, 222-229.

Sutherland, I. E. (1968). A Head-Mounted Three Dimensional Display*. Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, 757-764.

Symmes, D. L. (2006, November 20 2006). The Chopper. Accessed January 29 2014 at http://www.3dmovingpictures.com/chopper.html
Szirmay-Kalos, L., & Umenhoffer, T. (2008). Displacement Mapping on the GPU—State of the Art. *Computer Graphics Forum,* 27(6), 1567-1592.
Theile, G., & Plenge, G. (1977). Localization of Lateral Phantom-Sources. *Journal of the Audio Engineering Society,* 25(4), 196-200.
Tramberend, H. (1999). Avocado: A Distributed Virtual Reality Framework. *Proceedings of IEEE Virtual Reality 1999*, 14-21.
Treisman, A. (1996). The Binding Problem. *Current opinion in neurobiology,* 6(2), 171-178.
Troje, N. F., & Westhoff, C. (2006). The Inversion Effect in Biological Motion Perception: Evidence for a “Life Detector”? *Current Biology,* 16(8), 821-824.
Tschermak, A. (1931). Optischer Raumsinn. *Receptionsorgane II* (pp. 834-1000): Springer.
Ukai, K., & Howarth, P. A. (2008). Visual Fatigue Caused by Viewing Stereoscopic Motion Images: Background, Theories, and Observations. *Displays,* 29(2), 106-116.
Urbantschitsch, V. (1897). Über Störungen Des Gleichgewichtes Und Scheinbewegungen. [On disturbances of the equilibrium and illusory motions]. *Zeitschrift für Ohrenheilkunde,* 31, 234-294.
Valeriano, P. (1556). *Hieroglyphica*. Basel.
Van Gelder, T. (1995). What Might Cognition Be, If Not Computation? *The Journal of Philosophy,* 92(7), 345-381.
Varela, F. J., Thompson, E. T., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*: The MIT Press.
Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why Pictures Look Right When Viewed from the Wrong Place. *Nature neuroscience,* 8(10), 1401-1410.
Vishwanath, D., & Hibbard, P. B. (2013). Seeing in 3-D with Just One Eye Stereopsis without Binocular Vision. *Psychological science,* 24(9), 1673-1685.
Viswanathan, G., Buldyrev, S. V., Havlin, S., Da Luz, M., Raposo, E., & Stanley, H. E. (1999). Optimizing the Success of Random Searches. *Nature,* 401(6756), 911-914.
Von Helmholtz, H. (1866). *Handbuch Der Physiologischen Optik* (Vol. 9): Voss.
Voß, G., Behr, J., Reiners, D., & Roth, M. (2002). A Multi-Thread Safe Foundation for Scene Graphs and Its Extension to Clusters. *Proceedings of the Fourth Eurographics Workshop on Parallel Graphics and Visualization*, 33-37.
Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The Precedence Effect in Sound Localization. *The American Journal of Psychology,* 62(3), 315-336.
Walton, K. L. (1990). *Mimesis as Make-Believe : On the Foundations of the Representational Arts*. Cambridge, Mass.: Harvard University Press.
Wang, W., Jüttler, B., Zheng, D., & Liu, Y. (2008). Computation of Rotation Minimizing Frames. *ACM Transactions on Graphics (TOG),* 27(1), 2.
Warncke, H. (1941). Die Grundlagen Der Raumbezeuglichen Stereophonischen Uebertragung Im Tonfilm. [The fundamentals of room-related stereophonic reproduction in sound films]. *Ak. Zeitschrift,* 6(3), 174-188.
Warren, H. C. (1895). Sensations of Rotation. *Psychological Review,* 2(3), 273.
Waschbüsch, M., Cotting, D., Duller, M., & Gross, M. (2006). WinSGL: Software Genlocking for Cost-Effective Display Synchronization under Microsoft Windows. *Proceedings of the 6th Eurographics conference on Parallel Graphics and Visualization*, 111-118.
Wexler, M., & Droulez, J. (2003). Allocentric Perception of Space and Voluntary Head Movement. *Psychological Science,* 14, 340-346.
Wexler, M., Lamouret, I., & Droulez, J. (2001). The Stationarity Hypothesis: An Allocentric Criterion in Visual Perception. *Vision research,* 41(23), 3023-3037.
Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-Motion and the Perception of Stationary Objects. *Nature,* 409(6816), 85-88.

Wexler, M., & Van Boxtel, J. J. (2005). Depth Perception by the Active Observer. *Trends in cognitive sciences,* 9(9), 431-438.

Wheatstone, C. (1838). Contributions to the Physiology of Vision.--Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Philosophical transactions of the Royal Society of London,* 128, 371-394.

White, J. (1987). *The Birth and Rebirth of Pictorial Space*: Faber and Faber.

White, L. T. (1978). *Medieval Religion and Technology: Collected Essays*: University of California Press.

Whitney, W. D. (Ed.) (1895) The Century Dictionary, an Encyclopedic Lexicon of the English language (Vols. III). New York: The Century Co.

Wiederhold, B. K., Jang, D., Kaneda, M., Cabral, I., Lurie, Y., May, T., . . . Kim, S. (2001). An Investigation into Physiological Responses in Virtual Environments: An Objective Measurement of Presence. *Towards cyberpsychology: Mind, cognitions and society in the internet age*.

Wiggins, B. (2004). *An Investigation into the Real-Time Manipulation and Control of Three-Dimensional Sound Fields.* (PhD), University of Derby. Retrieved from http://derby.openrepository.com/derby/bitstream/10545/217795/1/BW_PhDThesis.pdf

Wilkes, J. (1810). Canicular Days. *Encyclopaedia Londinensis* (Vol. 3).

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3), 225-240.

Wolf, W. (2008). Is Aesthetic Illusion› Illusion Référentielle‹?› Immersion‹ in (Narrative) Representations and Its Relationship to Fictionality and Factuality. *Journal of Literary Theory,* 2(1), 99-126.

Wolf, W. (2013). Illusion (Aesthetic). In P. Hühn (Ed.), The Living Handbook of Narratology. Hamburg: Hamburg University Retrieved from http://www.lhn.uni-hamburg.de/article/illusion-aesthetic.

Wood, R. (1895). The 'Haunted Swing' Illusion. *Psychological Review,* 2(3), 277.

Woods, A. J. (2001, January 22). Optimal Usage of LCD Projectors for Polarized Stereoscopic Projection. *Photonics West 2001-Electronic Imaging*, VIII(5), San Jose, 5-7.

Woods, A. J. (2010). Understanding Crosstalk in Stereoscopic Displays. *3DSA Three-Dimensional Systems and Applications*, Tokyo, Japan.

Woods, A. J. (2012). Crosstalk in Stereoscopic Displays: A Review. *Journal of Electronic Imaging,* 21(4), 040902-040902.

Woods, A. J., Docherty, T., & Koch, R. (1993). Image Distortions in Stereoscopic Video Systems. *IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology*, 36-48.

Wright, M. (2005). Open Sound Control: An Enabling Technology for Musical Networking. *Organised Sound,* 10(03), 193-200.

Yaras, F., Kang, H., & Onural, L. (2010). State of the Art in Holographic Displays: A Survey. *Display Technology, Journal of,* 6(10), 443-454.

Yeh, Y.-Y., & Silverstein, L. D. (1990). Limits of Fusion and Depth Judgment in Stereoscopic Color Displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 32(1), 45-60.

Yu, J., McMillan, L., & Sturm, P. (2010). Multi-Perspective Modelling, Rendering and Imaging. *Computer Graphics Forum*, 29(1), 227-246.

Zahorik, P. (2001). Estimating Sound Source Distance with and without Vision. *Optometry & Vision Science,* 78(5), 270-275.

Zahorik, P. (2003). Auditory and Visual Distance Perception: The Proximity-Image Effect Revisited. *The Journal of the Acoustical Society of America,* 113, 2270.

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory Distance Perception in Humans: A Summary of Past and Present Research. *Acta Acustica united with Acustica,* 91(3), 409-420.

Zahorik, P., & Jenison, R. L. (1998). Presence as Being-in-the-World. *Presence: Teleoperators and virtual environments,* 7(1), 78-89.

Zhang, C., & Chen, T. (2004). A Survey on Image-Based Rendering—Representation, Sampling and Compression. *Signal Processing: Image Communication,* 19(1), 1-28.

Zhang, X., Hantke, K., Fischer, C., & Schröter, M. (2012). Performance of Polarization-Based Stereoscopy Screens. *3D Research,* 3(4), 1-5.

Zheng, J. Y., & Tsuji, S. (1990). Panoramic Representation of Scenes for Route Understanding. *Pattern Recognition, 1990. 10th International Conference on*, 1, 161-167.

Zheng, J. Y., & Tsuji, S. (1992). Panoramic Representation for Route Recognition by a Mobile Robot. *International Journal of Computer Vision,* 9(1), 55-76.

Zimmons, P., & Panter, A. (2003). The Influence of Rendering Quality on Presence and Task Performance in a Virtual Environment. *Proceedings of IEEE Virtual Reality 2003*, 293-294.

Zone, R. (2007). *Stereoscopic Cinema and the Origins of 3-D Film, 1838-1952*. Lexington: University Press of Kentucky.

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of the University of New South Wales' products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to *http://www.ieee.org/publications_standards/publications/rights/rights_link.html* to learn how to obtain a License from *RightsLink*.