

**Navigating 3d Scatter Plots in
Immersive Virtual Reality
Geoff Gray**

**A thesis submitted in partial
fulfillment of the requirements for
the degree of:
Master of Design**

**University of Washington
2016**

**Committee:
Axel Roesler
Tad Hirsch
Christian Marc Schmidt**

**Program authorized to offer degree:
School of Art + Art History + Design**

©Copyright 2016
Geoff Gray

University of Washington

Abstract

Navigating 3d Scatter Plots in Immersive Virtual Reality

Geoffrey E Gray

Chair of the Supervisory
Committee:
Dr. Axel Roesler
Chair, Interaction Design

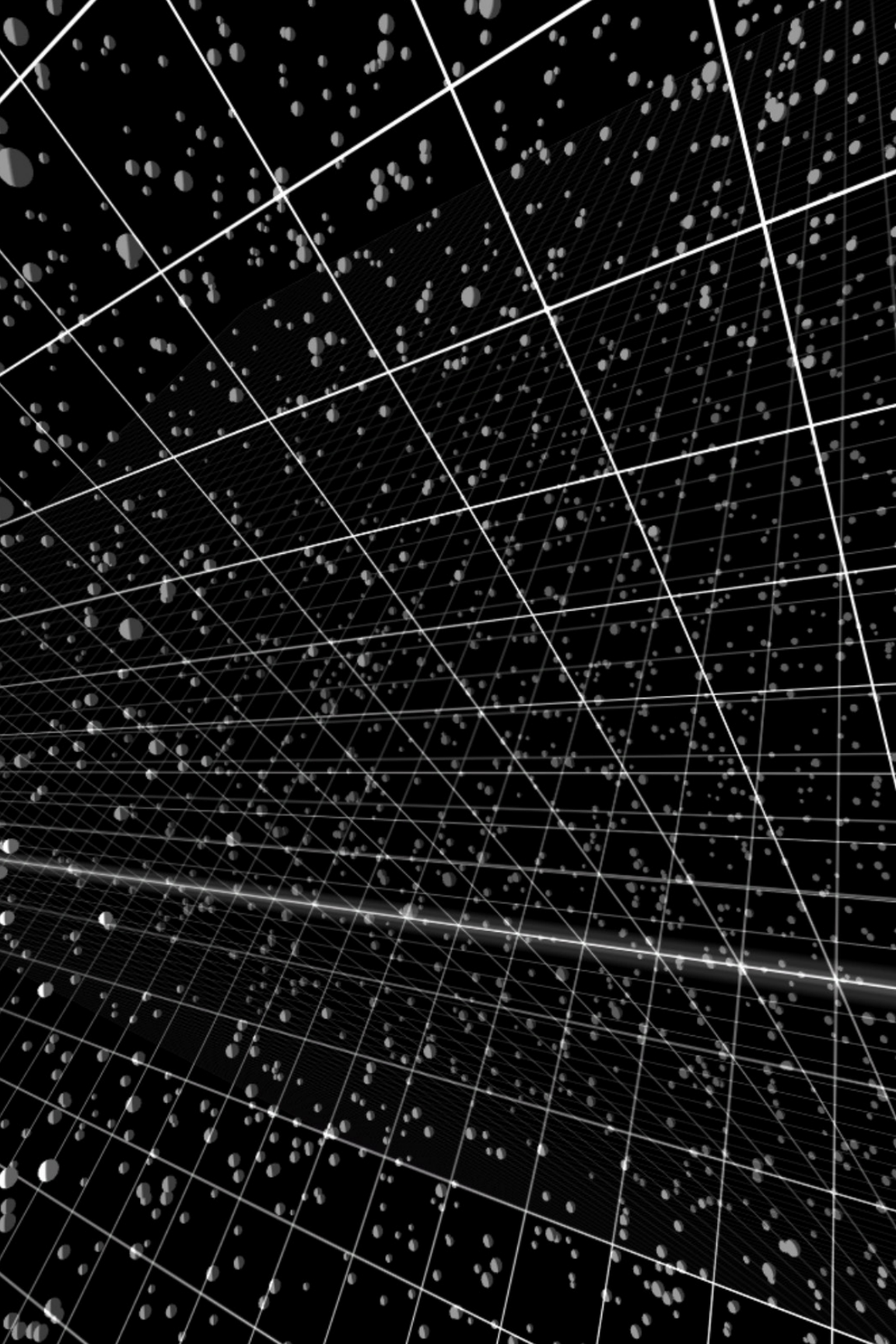
Virtual reality reveals spatially complex structures behind 3d data in an easily explorable way. The difficulty is in navigating this space while wearing a head-mounted display (HMD). One viable option for view control is a combination of a 3d mouse for camera location and head tracking of the HMD for the control of view direction. The scenes shown here display semi-randomized 3d scatter plots, each of which emphasizes a type of navigation that leads the viewer to discover structure behind the data. The three types of navigation are rotation, scale and alignment.

NAVIGATING
3D SCATTER
PLOTS IN
IMMERSIVE
VIRTUAL
REALITY

GEOFF GRAY

**NAVIGATING
3D SCATTER
PLOTS IN
IMMERSIVE
VIRTUAL
REALITY**

GEOFF GRAY



**Navigating 3d Scatter Plots in
Immersive Virtual Reality**
Geoff Gray

CONTENTS

PREFACE/ ACKNOWLEDGEMENTS	VII
INTRODUCTION	1
Abstract	1
Significance	3
BACKGROUND	5
Data Visualization	5
Virtual Reality	11
Related Work	12
PROCESS	23
Hardware	24
Software	26
PROJECT	29
Scene 1: Rotation	31
Scene 2: Scale	37
Scene 3: Alignment	43
Exhibit	53
CONCLUSION	59
BIBLIOGRAPHY	62

PREFACE

I must admit that working with Virtual Reality (VR) was never a consideration prior to my graduate studies. In all honesty, I'm still not entirely sure how I ended up here. Looking back, however, VR does satisfy many of my grad school "wants."

Foremost, VR is an emerging technology, mostly free from prescribed methods, techniques and applications. Nobody is entirely sure what to do with it. Sure, gaming is a given. It's just neat to fly a spaceship around a virtual cosmos. But I am not interested in givens. Surely there are some less-considered applications for this new technology.

One application—the one I have spent the most time with—is data visualization. How can VR add to the process of visualizing abstracted information? We experience non-virtual reality in three spatial dimensions. Surely, moving data visualization beyond the flat screens of a soon-to-be yesterday must have some immediate benefits. This book documents benefits, disadvantages and considerations behind data visualization—in this case 3d scatter plots—in virtual reality environments.

ACKNOWLEDGEMENTS

No thesis is ever one person's work. I would like to extend my utmost gratitude to a number of influential individuals.

First and foremost, I must say thank you to my family. Dad, I think you finally understand what I do! I hope to undo that soon. Mom, thanks for raising a curious rebel with a penchant for dry sarcasm. Have any of your other 30 children expressed interest in a particular field? (she raises chickens and affectionately refers to them as her "babies").

Thank you to a wonderful SigOth who's frequent words of encouragement kept my head up. Someday I will repay you!

In an order purely alphabetical, many thanks to my peers: Chad P. Hall, Jaewon Hwang, Catherine Lim, Joe Sparano and Erin Wilson. It has been a tough yet rewarding ride. Thanks as well to the to-be design masters with whom I shared studio space.

I also extend my gratitude to the faculty. Thank you Linda Norlen for helping me break the habit of excessive commas—Linda also helped this thesis take shape during the early stages. Tad Hirsh, thank you for the needed-but-not-wanted kick in the ass. I still wonder how insane (and hilarious) my thesis would have been with you as chair.

Additionally, Christian Marc Schmidt is absolutely deserving of notice. While we didn't get to work all that much on the thesis itself, I appreciate greatly our discussions.

Lastly there is Axel Roesler, the chair of my thesis committee. It has been an absolute pleasure and honor to work with you on this project. Without your support (and criticism!), this project simply would not be. I hope that everything I have learned here will be of some use to your future students.



Introduction

ABSTRACT (BRIEF)

Virtual reality reveals spatially complex structures behind 3d data in an easily explorable way. The difficulty is in navigating this space while wearing a head-mounted display (HMD). One viable option for view control is a combination of a 3d mouse for camera location and head tracking of the HMD for the control of view direction. The scenes shown here display semi-randomized 3d scatter plots, each of which emphasizes a type of navigation that leads the viewer to discover structure behind the data. The three types of navigation are rotation, scale and alignment.

ABSTRACT (EXTENDED)

Virtual reality reveals spatially complex structures behind 3d data in an easily explorable way. The difficulty is in navigating this space while wearing a head-mounted display (HMD). One viable option for view control is a combination of a 3d mouse for camera location and head tracking of the HMD for the control of view direction. The scenes shown here display semi-randomized 3d scatter plots, each of which emphasizes a type of navigation that leads the viewer to discover structure behind the data. The three types of navigation are rotation, scale and alignment.

While the 3d mouse allows the viewer to "get up and fly around" the virtual environment, the HMD simultaneously allows changes in visual display through minor adjustments of the head in ways that 3d projections on 2d screens can not. These two types of vision—vision being inherently linked to movement—are referred to as ambulatory and ambient vision, respectively.

Visual cues such as illumination, shadows, textures, transparency/opacity, occlusion, the inclusion of a ground plane and, most importantly, motion parallax all improve the sense of immersion, feeling physically present in a non-physical yet plausible environment.



2 Immersive displays are not a new idea. This is Hugo Gernsback: "The Father of Science Fiction" wearing his invention, TV Glasses in 1908.

SIGNIFICANCE

Data visualization, communicating abstract information through graphic display, is not new. The ability to interact with data, however, is relatively recent. As software efficiency and consumer-level computer hardware improve, sophisticated data visualizations are becoming computationally trivial.

In the past, the practice of data visualization would have been the domain of specialist statisticians, engineers and academics.

Today's connected world has placed data visualization into the mainstream. From business analysts to journalists, a growing number of professions are producing data visualizations in their regular work. New tools—such as d3 (Data-Driven Documents), IBM's Many Eyes and Google Charts—have lowered the barrier to entry for producing data visualizations on the web by providing templates for common data visualization tasks. There is, however, a rapidly emerging display technology that may change the ways in which data is viewed and experienced.

A number of virtual reality (VR) and augmented / mixed reality (AR / MR) headsets will be released this year, 2016. Companies such as Oculus, HTC, Sony and Microsoft are devoting considerable resources to the development of these headsets. Like many emerging technologies, content developers are producing wildly entertaining, immersive and sometimes disorienting experiences. Nearly all of this content is intended for entertainment purposes. However, these technologies, as they become commonplace in the living room, will likely make their way into the office. So, what are potential

applications of these devices in the realm of data visualization?

One immediate benefit is the addition of another dimension to data display. Quite obviously, being able to view a 3d data visualization with a 3d display is logical. The benefits, though somewhat obvious, are not inconsequential. We experience the world in three dimensions, we should experience data in three dimensions.

Using 3d scatter plots as a logical launching board, this project seeks to explore potential applications of data visualization for VR devices. In particular, what design and visual principles are important in developing display and navigation techniques?



Background

DATA VISUALIZATION

Data-driven science is revolutionizing the breadth and depth of human knowledge and discovery. Over the past few decades, the world has witnessed a deluge of data, an avalanche of information. Legions of petite yet powerful sensors are being deployed around the globe, fueling studies on climate change and furthering findings in the atmospheric sciences. Contemporary scientific instruments—from subterranean particle accelerators spanning nations to space telescope satellites orbiting the Earth—are offering confirmatory evidence to physicists' theories of the origins of space and time. Simulations performed on the grandest, most powerful supercomputers are modeling the behavior of matter at sub-atomic levels. Even our own activities, especially those done on electronic devices, are collectively revealing insights into the whole of contemporary human behavior.

For science, the great challenge presented by the data deluge is making sense of and using effectively these expanding corpora of information. Exploiting the high bandwidth channel of human visual perception, scientific visualizations are critical in enabling people to understand, explain, and explore vast amounts of data from a myriad of sources. By assisting the exploration of hypothetical situations through fine control of variables and perspectives, visualization is poised as an integral avenue

toward scientific discovery. Effective visualization, especially when coupled with the potential of customization, encourages researchers to compare alternative models and data sets in order to make connections or inferences between what would otherwise be incomprehensible—raw data. Although scientific visualization is still a relative newcomer to the realm of academic discourse, frameworks for understanding how users make sense of and gain insight from visualization have seen some early attempts. This paper surveys these frameworks and maps them to contemporary examples of effective visualizations within the sciences in order to better understand how discoveries are made through visualizations.

Like language, visualization is a tool to present information and exists in many forms. Visualization can be seen as the parental category of its closely related and often overlapping sub-fields including educational visualization, information visualization, scientific visualization, statistical graphics, data visualization, molecular graphics, isosurface reconstruction and thematic cartography, to name a few. Traditionally, scientific visualization—an interdisciplinary field—has been concerned with realistic renderings of naturally occurring phenomena in areas of biology, medicine, engineering, architecture, climate, etc. Scientific visualizations often include three spatial dimensions and may include a

temporal dimension. Data visualization, representing information that has some level of abstraction, differs from scientific visualization in that the dimensions of data often extend beyond our familiar three (with or without time), making the task of mapping the data to two- or three-dimensional domains nontrivial.

FROM DATA TO KNOWLEDGE

The terms data, information, and knowledge appear often in the visualization community. For the most part, they describe varying levels of abstraction, understanding, or truthfulness. Alternatively, these three terms can refer to types of input as in “data visualization,” “information visualization,” and “knowledge visualization.” Here, anything being represented visually is considered data. However, it is still helpful to introduce these concepts as appropriate to the first scheme. Data is the lowest level of abstraction. It represents objective facts or observations (which may not necessarily be true) that are unorganized and therefore mostly meaningless to people. Information is data that has been captured and organized in some meaningful way that makes it useful to people—put rather simply, information is that which informs—and is often an attempt to clarify uncertainty. Knowledge, a concept more difficult to define, is acquired by understanding, making sense of, and extracting meaning from information. As an example, consider commuting to work or school. Most certainly, there are a multitude of routes to work and each of these routes all span a certain distance. These distances, being objective facts, will continue to exist whether or not they are measured and in this sense are raw data. Information, following

this analogy, would be a time estimate of each particular route. This estimate is composed of a different types of data such as traffic, road construction, and accidents along the route. Knowledge, here, would be a summation of these traffic reports when placed in context: on your way home the day before perhaps you noticed that there would be a large football game the following day so you decide to take a route further away from the stadium. Mapping visualization to this analogy, consider getting route suggestions from Google Maps for the first few days or weeks after starting a new job or school program. Google Maps provides (mostly) accurate estimates for a small selection of the most efficient routes to your destination. After a few rounds of taking different routes at different times, you notice a pattern: one route seems to speed up drastically after 5:45pm. Forming knowledge from this information, you adjust your schedule so that you leave a few minutes before 5:45pm. You have gained knowledge about your ideal route home by noticing patterns of a visualization.

Exploiting our innate human ability to make spatial inferences and decisions, visualizations map elements and relations from abstract domains onto concrete domains allowing a recursive transference of spatial inferences back to abstract domains [5]. However, human capabilities of both visual and cognitive information processing have limits and are, by nature, systematically biased. Designing effective visualizations requires an awareness of these constraints. Methods to address these limits and biases include selecting and focusing on essential information points, reducing visual clutter (noise), and translating ideas that are not inherently visual into visual domains. These

methods of working with human constraints reveal visualization as an explicitly human endeavor. While computer science, particularly the field of artificial intelligence, works towards replacing human judgment with automation, visualizations are purposely designed to include the human by extending our capabilities [5]. In effective visualizations, user interaction is often both expected and configurable. The route from exploration to discovery through interaction proceeds in many complex ways.

THE CHALLENGE OF “BIG” DATA

“Big” data is relative. In the sciences, big data has been used to describe data sets large enough to warrant time on supercomputers. Closer to the realm of business, big data has been described by the “3vs” model as having high volume, high velocity, and/or high variety information. Here volume refers to the size and quantity of data, both breadth and depth. In the age of information, where data can be seen as an asset, organizations are often reluctant to discard data which in turn leads to increased volumes of data. Velocity, the speed of data input, processing, and output, is especially important for data processing that occurs in near real time environments such as client–server interactions. Data variety is often the result of incompatible formats, non–aligned structures, and inconsistent semantics. Often, these data sets require technological advances in hardware and/or software to be processed appropriately and, therefore, to be of increased utility. For big data in the sciences, one especially problematic characteristic is its networked nature, a consequence of increased capabilities in mining and aggregating data, particularly through the web. Issues along the data supply chain

that are unnoticed or unaccounted for will most likely lead to undesired or inaccurate results, apropos of the “garbage in–garbage out” principle of computer science.

Since data used for scientific visualizations usually involves some form of human intervention, the data’s provenance–descriptions of its origins and the processes/modifications the data has undergone–should be appropriately annotated. Large volumes and complex analyses are common in scientific exploration. Annotation, through notes and/or visualizations of the processes applied to the data over time, is important both in preventing duplication of efforts as well as building upon the discoveries of others. Additionally, provenance retains documentation that is important to preserving data, determining its quality or veracity, tracking its authorship, reproducing and validating its results, and navigating processes of reasoning forward and backward. One tool for visualizing provenance is VisTrails, an open source platform that captures modifications applied to other data visualizations and can be integrated with interactive tools that, previously, could not easily be captured in a workflow system.

Big data is also a moving target. The previously overwhelming “big” data of yesteryear is rapidly dwarfed by the exponential growth of our capacity to generate and, to a lesser degree, store data. For example, in 1999 the total volume of information generated that year was estimated to be two exabytes (two billion gigabytes). To put that into perspective, the same study estimated that all words ever spoken by human beings would require five exabytes of storage. A Cisco paper recently reported that in 2013, internet traffic alone was measured at 1.68 exabytes per

day. While most of this traffic was generated by consumers, scientific studies are also generating extraordinary volumes of data. The particle detector, a measurement device at the Large Hadron Collider in Switzerland, even after rejecting 199,999 of every 200,000 particle collisions, produced 13 petabytes (13 million gigabytes) of data in 2010. In addition to measuring observable phenomena, scientific data is also generated through simulations. In 2002, Lawrence Livermore National Laboratory performed hydrodynamic instability simulations generating tens of terabytes of data. While that volume might seem small in today's numbers, the simulations were performed two years later in real time on the Blue Gene/L installation at the Supercomputing 2004 exhibit. At that time, the Blue Gene/L was the fastest supercomputer in the world. The data from these simulations required visualization and analysis to support and affirm its underlying model and to glean new perspectives into its fundamental physics. Advances in both visualization and data analysis techniques were required to allow performance of simulations in real time during execution.

GAINING INSIGHTS THROUGH VISUALIZATION

Although visualization involves advanced and often complex transformations, mappings, and interaction techniques, its ultimate goal is to extend human capability through visual perception in order to assist identification of trends, patterns, connections, and abnormalities in data sets. By better understanding how insight is gained through these processes, visualization practitioners will be better equipped to capitalize on the strengths of human visual

perception and increase the likelihood of novel insights and discoveries.

Within the visualization community, there have been a variety of attempts to define insight. Similar to the concept of big data, insight may be better captured through an array of characteristics rather than a common agreed upon, concrete definition. One scheme of insight characteristics that has seen some level of adoption is offered by North. He describes insights as having the following characteristics: complex, deep, qualitative, unexpected, and relevant. Insight is complex. It synthesizes large amounts of data, not just single data points. Insight is deep. It aggregates and builds upon itself over time, often leading to further questions and, therefore, potentially further insight. Insight is qualitative. It is not exact and can be uncertain or subjective with multiple levels of resolution. Insight is unexpected. It is unpredictable and serendipitous. In Feynman's words, "...what is not surrounded by uncertainty cannot be the truth." Insight is relevant. It aims to connect data to existing knowledge, giving data meaning. As the last step in the process of data to discovery via visualization, understanding insight—in particular how people gain insight from visualizations—is critical.

In order to better understand the process of gaining insights, a framework of sense making through visualization is helpful. Similar to the concepts of big data and insight, sense making through visualization can also move towards a definition through a scheme of characteristics: iterative and cyclical, a discovery and creative process, and lastly retrospective. Sense making through visualization is iterative and cyclical, a process of data collection and representation

“The purpose of visualization is insight, not pictures.”*

adjustments. Sense making through visualization is at once a discovery and a creative process, involving interpretation and invention. Sense making through visualization is retrospective in that often people construct a framework before collecting information. If the information fits with the established framework, the framework is confirmed; if not, the framework is discarded or updated in order to explain the information. This scheme for sense making places insight as an intermediary for discovery in that insight can reveal the need for a framework to explain information observed.

As previously explored, insight (and therefore discovery) gained from visualization does not usually result from initial impressions. Rather, insight is gained through a series of processes. Through a survey of literature on “InfoVis” (Information Visualization) Yi et al. identified four processes through which visualization users gain insight: Provide Overview, Adjust,

Detect Pattern, and Match Mental Model.

Provide Overview can be described as introducing the data in “big picture” format. While an overview of data sets does not necessarily yield direct insight, it assists people in establishing the framework in which to place information (data) observed. From here, the user is able to spot particular areas for further investigation and exploration of the data sets. Provide Overview allows people to review what they already know and reveals the extent to which they could gain new knowledge from the data sets under observation. Provide Overview can reveal the overall structure of complex systems and processes.

Adjust implies flexibility for customization of the presentation of data. Visualization users are able to view/hide certain parameters, adjust the level of abstraction, and narrow or expand the range of selection. Adjust implicitly assumes some level of interaction. When exploring large amounts of data sets, visualizing them in their

* Card, Mackinlay, and Schneiderman, 1999.

entirety is often overwhelming, riddled with noise. By selecting prescribed ranges and/or types of data, the user is able to observe only particular data of interest. By allowing comparison of a variety of data types and measurements, previously unnoticed patterns or relationships can be revealed. One common technique is the inclusion of “sliders” in which the display instantly responds to user-input parameters. Additionally, grouping is an effective technique to abstract large data sets into more manageable pieces, reducing the time the user spends searching. This allows the user to focus their attention on a more manageable subset of the data.

The Detect Pattern process allows the user to not only confirm established hypotheses, models, or frameworks but also to spot the unexpected. Finding particular distributions, trends, frequencies, or structures in the data sets are included in the category of unexpected and can often be the basis of a new insight.

Match Mental Model means providing a visually digestible representation of data. Often, this reduces the cognitive load of understanding the data and extends the user’s ability to recognize familiar occurrences, closing the distance between data and the user’s real-world understanding of certain phenomena. For example, both wind and ocean currents, at any particular point in time, are not visible phenomena. However, our mental model of wind represents it as movement over time. From here, even static (lacking a temporal dimension) vector fields can assist human understanding of air or water currents. Other techniques include visualizations of radioactivity or

electromagnetic fields through hue and/or value visual metaphors.

CONCLUSION

The desire to explain the world in which we live—from fleeting subatomic particles to the molecular building blocks of life to supernova explosions that forged the elements of which we are composed—is an inherently human pursuit that has been served quite well by scientific endeavors. Visualizations prove to be powerful extensions of human intelligence, yielding knowledge that could not be achieved through computation or direct observation alone. Amplifying our immense potential for insight, effective visualizations facilitate the generation, evaluation, confirmation, rejection, and exploration of the hypotheses of our studies. By studying how researchers in data-driven science gain insights from visualization tools, visualization practitioners will be better equipped to produce tools for scientific discovery.

VIRTUAL REALITY

The term “virtual reality” (VR) is quite ambiguous. While it once could have been interpreted as something that only occurs in only in science fiction—yes, the Matrix...—virtual reality is quickly entering the household lexicon. More specifically, virtual reality is commonly accepted as a way to describe head-mounted displays (HMD) with stereoscopic display. There are two screens, one for the left eye and one for the right. This is referred to as a stereoscopic display. The advantage of this is a convincing illusion of depth along with an immersive field of view.

Of course, VR can be accomplished in other ways. I would argue that a blindfolded taste- or smell-test is a virtual reality, in a sense. One or more senses have been “blinded” to outside influence. For the purposes of this thesis, VR usually refers to a head-mounted display (HMD), in this case an Oculus Rift Development Kit 2 (DK2). The consumer version of the Rift is soon to be available which includes stereo headphones and an updated display.

VR is not a new idea. Like many technologies, science fiction explored applications of this technology well ahead of its practical availability. Perhaps the earliest functional virtual reality, as the term implies today, dates to 1987 by VPL Research, a company run by Jaron Lanier. VPL Research owned a number of patents for VR technologies.

One seminal development for VR is in Jonathan Waldern’s demonstration of “Virtuality” at a Computer Graphics exhibition. His product was an arcade machine that “immersed” players.

The 1990s had a handful of unsuccessful VR devices. In terms of entertainment, Sega VR was

largely a flop. Researchers at the University of Illinois at Chicago, under the Electronic Visualization Laboratory, produce a multi-projector room that became CAVE: Cave Automatic Virtual Environment. CAVE was used in a number of data visualization applications.

One of the first consumer VR devices, the Virtual Boy by Nintendo, was released in 1995. The device had underwhelming sales and was subsequently discontinued. This was my first interaction with VR. I was far from sold.

Fast-forward to 2010, consumer-grade cell phone displays are pushing the resolution boundaries of LCD screens. 72 pixels-per-inch (PPI) is history. We are now seeing displays upwards of 400 PPI. Along with this comes new potential for VR headsets, much more affordable than in the past.

Palmer Luckey, a lifelong VR advocate, founded Oculus VR in 2010. His technology was the first to reach a 90-degree field of vision. Later purchased in 2014 by Facebook for approximately \$2 billion, Oculus is soon to release a consumer version of the Rift. I have been working with a Developer Kit 2 (DK2).

RELATED WORK

Nearly all of the related projects I've seen have come from universities around the world. Given the technological challenges pertaining to VR—both hardware and software—it is not surprising that most of these projects were done by computer science or computer graphics experts. I have only seen a handful of documented VR data visualization projects from the private sector. One obvious explanation for this could be intellectual property rights.

The following are some of these projects. I have provided the abstract (if any) and my own brief summary of each project, along with images and image descriptions as written by the authors.

Immersive and Collaborative Data
Visualization Using Virtual Reality
Platforms

Ciro Donalek, S. G. Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, Andrew Drake
California Institute of Technology

Scott Davidoff, Jeffrey S. Norris
Jet Propulsion Laboratory

Giuseppe Longo
University Federico II

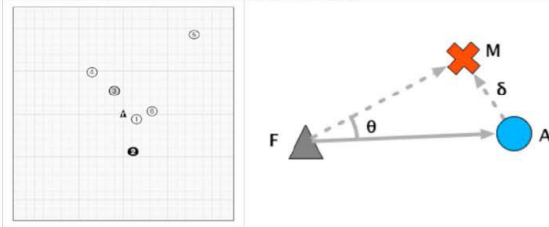
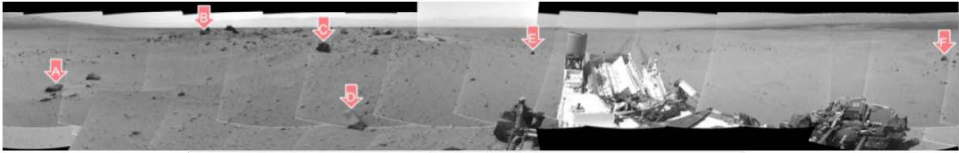
Abstract – Effective data visualization is a key part of the discovery process in the era of “big data”. It is the bridge between the quantitative content of the data and human intuition, and thus an essential component of the scientific path from data into knowledge and understanding. Visualization is also essential in the data mining process, directing the choice of the applicable algorithms, and in helping to identify and remove bad data from the analysis. However, a high complexity or a high dimensionality of modern data sets represents a critical obstacle. How do we visualize interesting structures and patterns that may exist in hyper-dimensional data spaces? A better understanding of how we can perceive and interact with multidimensional information poses some deep questions in the field of cognition technology and human-computer interaction. To this effect, we are exploring the use of immersive virtual reality platforms for scientific data visualization, both as software and inexpensive commodity hardware. These potentially powerful and innovative tools for multi-dimensional data visualization can also provide an easy and natural path to a collaborative data visualization

and exploration, where scientists can interact with their data and their colleagues in the same visual space. Immersion provides benefits beyond the traditional “desktop” visualization tools: it leads to a demonstrably better perception of a datascape geometry, more intuitive data understanding, and a better retention of the perceived relationships in the data.

–

This project works with a number of development techniques and tools to explore high-dimensional data sets. High-dimensional data visualization remains a challenge as humans can realistically only discern upwards of 10 to 15 visual dimensions of data (size, shape, position, color, etc.).

One key benefit to immersive VR, the authors argue, is collaboration. Multiple people can be “present” in the same data set while each having a unique view.



Top: the panoramic mosaic of the Curiosity Mars rover images used in this experiment, with some of the features of interest indicated with the arrows. Bottom left: an example of a map with the estimated relative positions of different features. Bottom right: a schematic illustration of the geometry used in the computation of distance and angle errors.

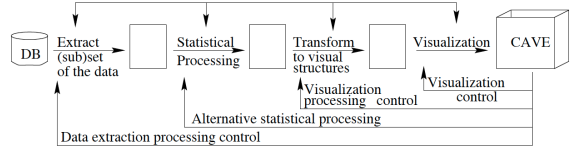
Methods for Visual Mining of Data in Virtual Reality

Henrik R. Nagel, Erik Granum, and Peter Musaeus

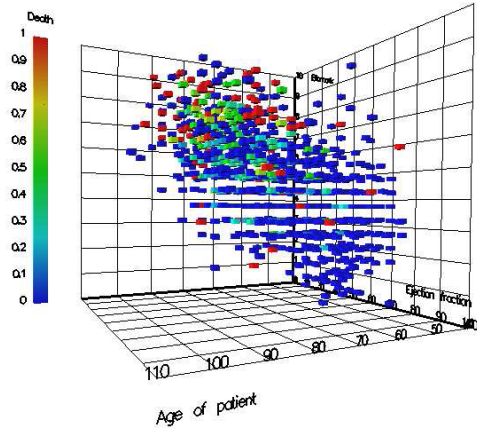
Lab. of Computer Vision and Media Technology, Aalborg University, Denmark

Abstract – Recent advances in technology have made it possible to use 3-D Virtual Reality for Visual Data Mining. This paper presents a modular system architecture with a series of tools for explorative analysis of large data sets in Virtual Reality. A 3-D Scatter Plot tool is extended to become an “Object Property Space”, where data records are visualized as objects with as many statistical variables as possible represented as object properties like shape, color, etc. A working hypothesis is that the free and real-time navigation of the observer in the immersive virtual space will support the chances of finding interesting data structures and relationships. The system is now ready to be used for experiments to validate the hypothesis.

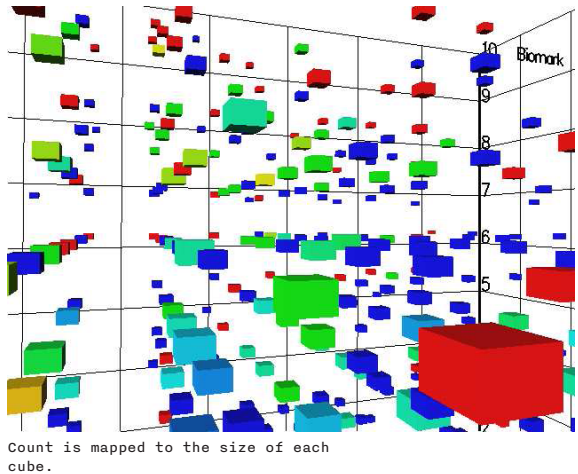
This project documents the development of a number of visualizations intended to be used as a testing ground for a viewer's capacity to recognize data structures and relationships. Some of these structures are patterning, clustering, outliers and correlation. Parameters, or visual cues, are position, pose, size, shape, color and texture. The authors indicate that encoding of a larger number of variables visually within the field of visual data mining are promising. I have been unable to find a follow up this publication.

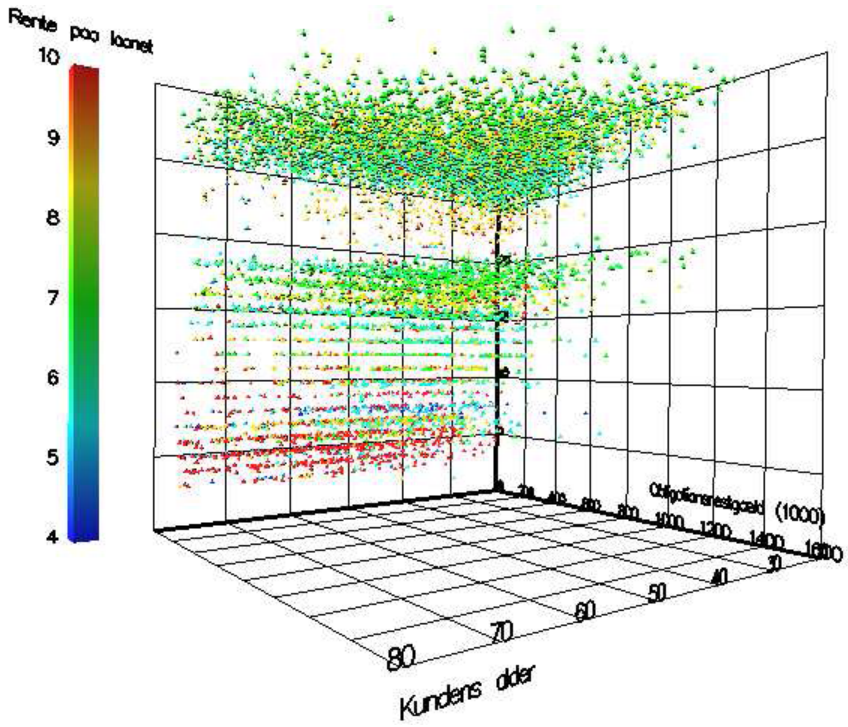


Data Flow and Interaction Patterns

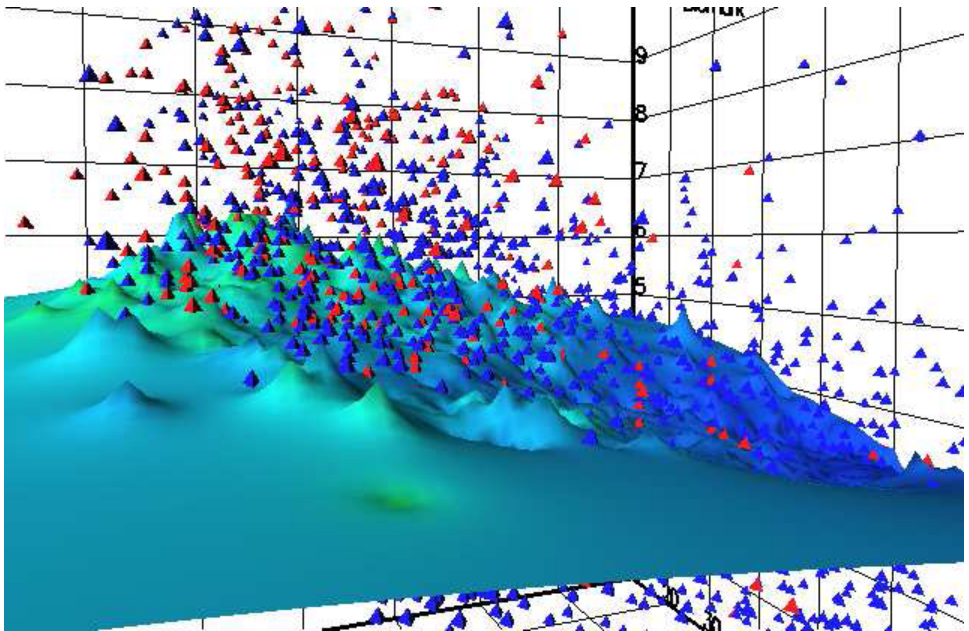


Visualizing 4 statistical variables.

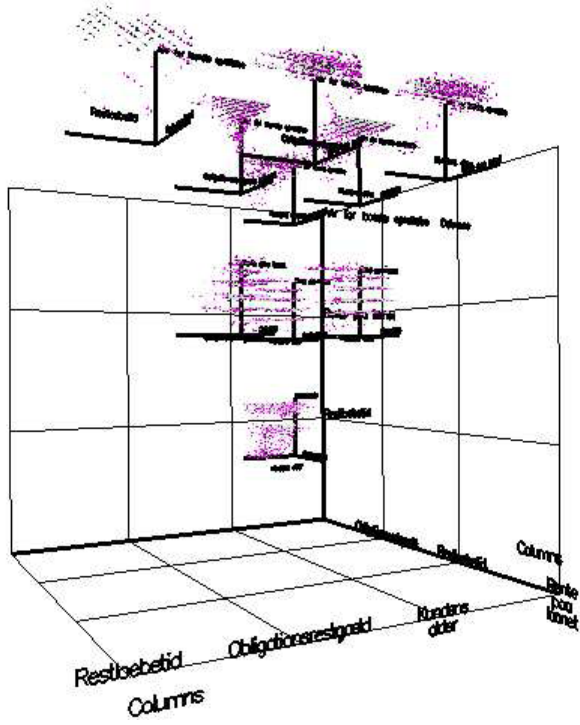




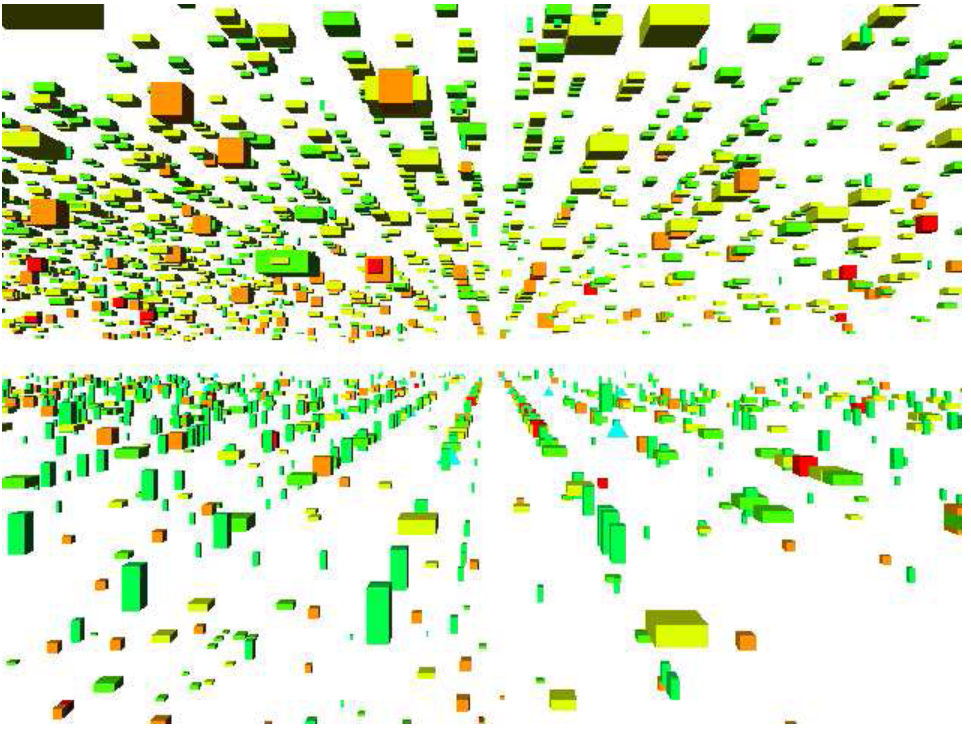
Scatter Plot.



Scatter Plot Matrix.



Scatter Plot with surface.



A look inside Object Property Space, using position, color, shape, and size to represent statistical variables.

Interactive exploration of immersive illuminated 3D scatterplots
ProjINF

Rafael Jarocsh, Sebastian Grund, and Sascha Sprött

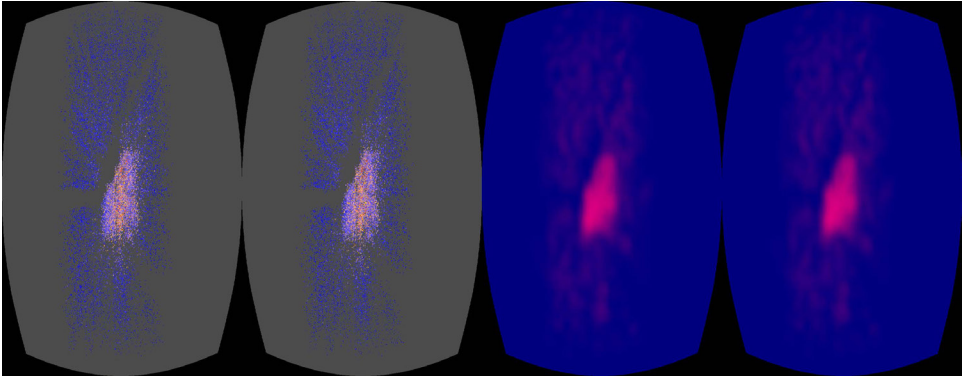
Abstract – The analysis of data is very significant in modern times. Over time, the amount of data grows and digital devices become increasingly important. To ensure progress and to understand the relationship between data, data-mining should be utilised, as it is now an indispensable tool. Data is usually represented in two dimensions. It is well known that 3D representations, such as Space-Time-Cubes, can be problematic on common 2D Desktops. Overlapping of data and the absence of depth perception are only a few of the many problems associated with 3D information visualisation. With the development of the Oculus Rift, a head-mounted-display, comes a multitude of possibilities. In the ability to track head movements and simulate depth perception, the Oculus Rift opens up new worlds in the field of data-mining. With this new technology, analysing three dimensional data seems to be more feasible.

–

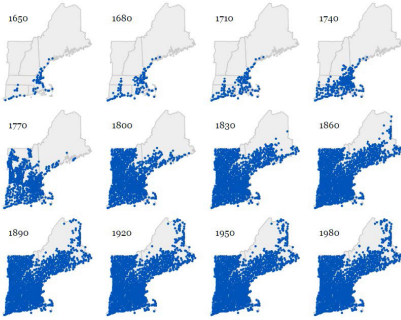
This project is perhaps closest to my own work. Some interesting differences are the authors' descriptions of four types of immersion, though they cite Ernst W. Adams, Staffan Bjork and Jussi Holopainen as the originators of these descriptions. The four types of immersion are: tactical immersion where higher brain functions are mostly shut down; strategic immersion which is commonly experienced by chess players; narrative immersion where the reader/viewer is engrossed in a

story; and spacial immersion where the simulated world is convincingly "real." Also of interest here is the functionality of a "selection cube," similar to a draggable mouse cursor but in 3d.

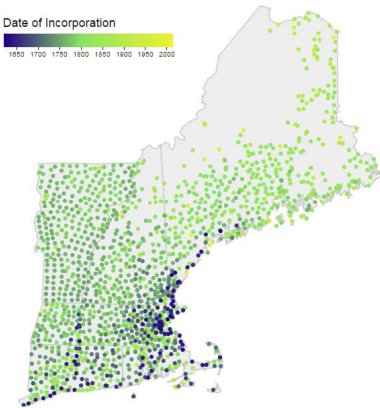
The authors conclude that VR is uniquely suited for geographical or geospatial data visualizations. I would extend this to encompass any data set with three spatial dimensions and, possibly, one temporal dimension.



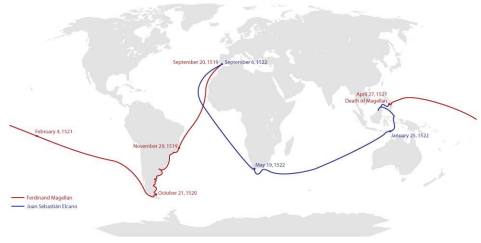
Scatter Plot - Vast Challenge.



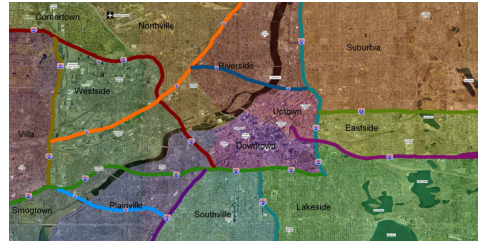
Multiple maps.



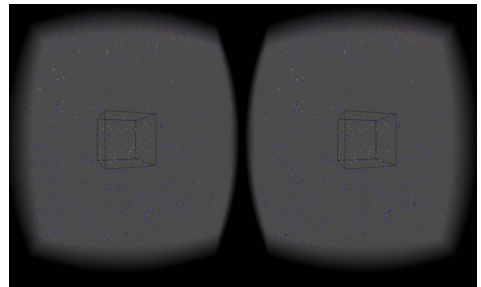
Static Maps. Time as map data.



Static Maps. Time as line on map.



Vast Challenge 2011.



Selection Cube.



Process

This section describes the hardware and software used in the project. The Oculus Rift DK2 has fairly specific hardware requirements as rendering is far more computationally demanding in stereoscopic displays, roughly four times that of a traditional, 2d display.

As for design principles, I rely primarily on *The Ecological Approach to Visual Perception* by James J. Gibson. This seminal—and groundbreaking—book looks to the natural world, ecologies, as a foundation for explaining visual perception systems. While the book is quite dense, I have identified a number of design considerations that are applicable to designing for VR. These are described more extensively in the Project section.

VR is notorious for its capacity to induce motion sickness. This is mostly due to visual senses receiving input that is in conflict with vestibular input. The effect is similar to reading while in a car. There are, however, a few design principles that can reduce the likelihood of motion sickness. Some of the most important considerations are:

- Avoid high-contrast flickering.
- Use real-world scales.
- No objects take up majority of user's view.
- No floating in a void.
- Provide static reference points.
- Points of Interest are in center of view.
- Sufficient frame rate & latency.

In the first weeks of this project, I did experience minor motion sickness a number of times. Now that I have spent considerable hours wearing and working with the device, it seems I have acclimated. This is a known phenomenon.

HARDWARE

These are specifications for the computer used in this project. It runs quite smoothly, even when rendering to both the Oculus Rift via HDMI and to a monitor via Display Port.

One important note is that the Oculus Rift DK2 did not work with USB 3.0 or USB 3.1. Rather, I had to use an external USB 2.0 hub.

I used a 3Dconnexion Space-Mouse for control of camera location, both in the project and in it's development.

CPU

Intel Core i5-6600K 3.5GHz
Quad-Core Processor

CPU Cooler

Corsair H100i GTX 70.7 CFM Liquid
CPU Cooler

Motherboard

Asus MAXIMUS VIII GENE Micro ATX
LGA1151 Motherboard

Memory

G.Skill Ripjaws V Series 16GB (2 x
8GB) DDR4-2400 Memory

Storage

Samsung 850 EVO-Series 250GB
2.5" Solid State Drive
Seagate Barracuda 2TB 3.5"
7200RPM Internal Hard Drive

Video Card

MSI GeForce GTX 970 4GB Twin
Frozr V Video Card

Case

Corsair 350D Window MicroATX Mid
Tower Case

Power Supply

EVGA SuperNOVA GS 650W 80+ Gold
Certified Fully-Modular ATX Power
Supply

Optical Drive

Asus DRW-24B1ST/BLK/B/AS DVD/
CD Writer

Operating System

Microsoft Windows 10 Pro OEM
64-bit

Wireless Network Adapter

TP-Link TL-WDN4800 PCI-Express
x1 802.11a/b/g/n Wi-Fi Adapter

Monitor

BenQ GW2765HT 27.0" 60Hz
Monitor



Oculus Rift DK2.



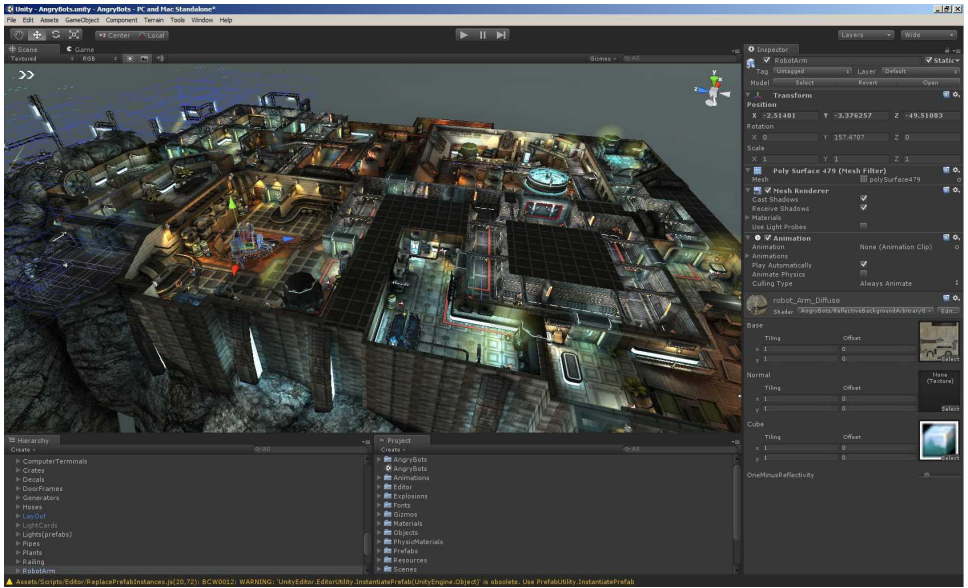
3Dconnexion SpaceMouse.

SOFTWARE

Unity, a game development engine, was the primary software used for this project. The learning curve is quite high but I would highly recommend Unity. Integration with the Oculus Rift DK2 was rather seamless. The VR plugin for Unity is highly functional, nearly plug and play. I used Unity 5.3.4 with an educational license. Unity also provides a free version which, it seems, has nearly all of the features as the Professional or Educational version. Unity appears to be the software of choice for VR development. Other options are Unreal, another game engine, and pure C++ through the Unity SDK.

Some level of programming is necessary for any Unity project. Unity's MonoDevelop is a fantastic script editor. However, I chose to work with Visual Studio. C# (C Sharp) was my language of choice though Unity also supports JavaScript. I chose C# as documentation for this language is far more extensive than for JavaScript (Unity's documentation is fantastic).

The SpaceMouse does have a plugin for Unity though documentation is nonexistent. I have included the scripts used in each scene in the Project section.



The Unity engine interface.



Project

The core of this project resides in three “scenes.” Each scene demonstrates a type of navigation. Each type of navigation leads the viewer to discover a type of structure behind the data set. The three scenes are:

SCENE 1: ROTATION

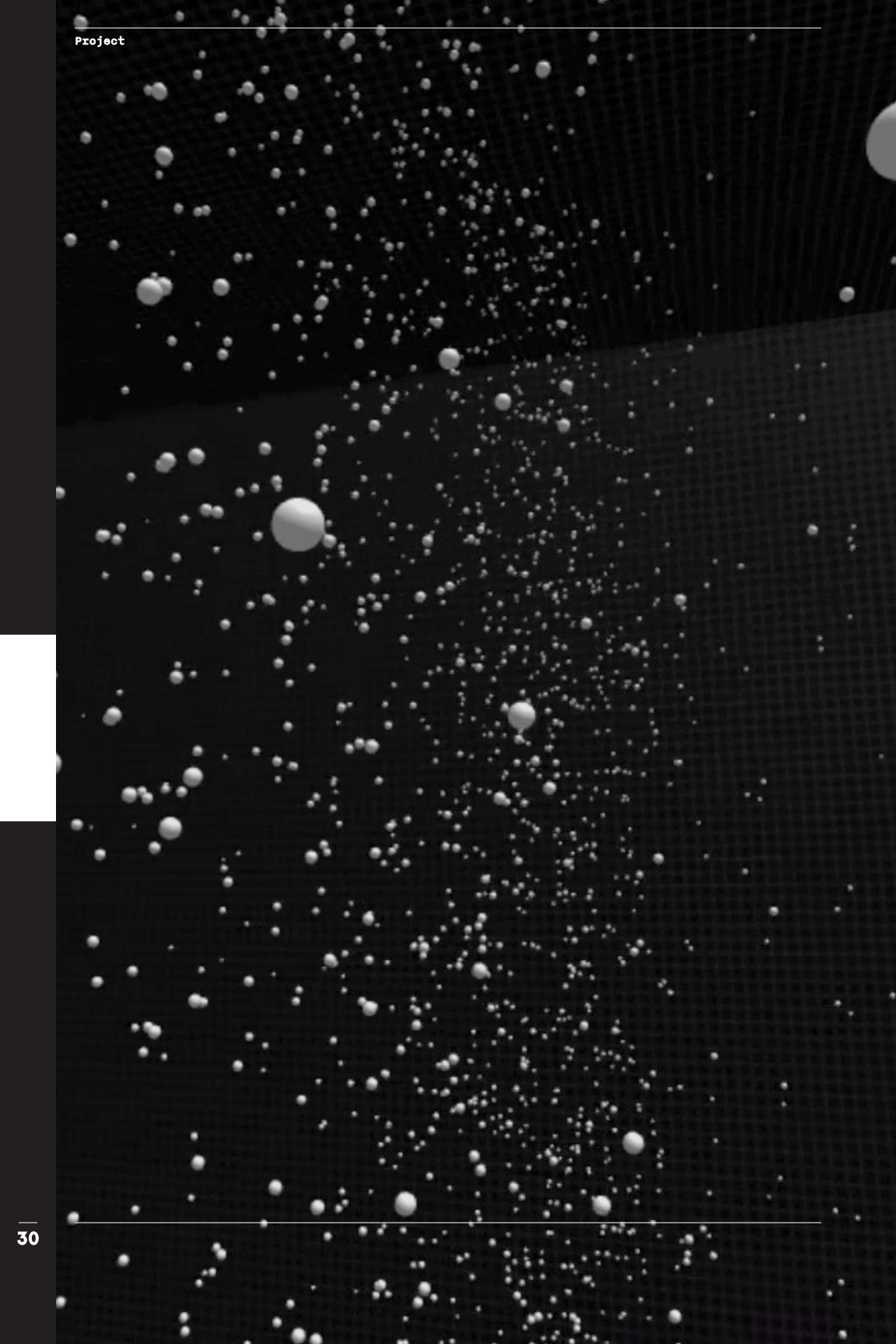
SCENE 2: SCALE

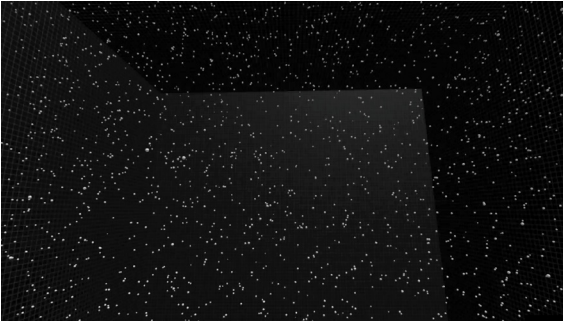
SCENE 3: ALIGNMENT

Naturally, documenting 3d scenes in 2d has its limitations. The scenes are intended to be viewed in person. This is why the construction of an interactive exhibit was crucial to this thesis.

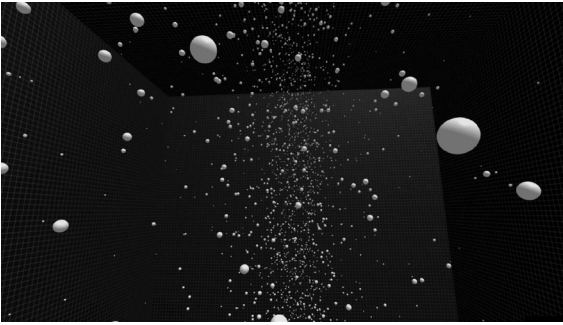
In this section, I showcase a initial, starting view in which one or more aspect of the data set’s structure is mostly indiscernible and an ending view in which the structure behind the data is revealed. This allows the viewer to have a mental model of the data set.

In addition, I have also included a timeline view of the scene. These thumbnail images show, roughly, user interaction over time.

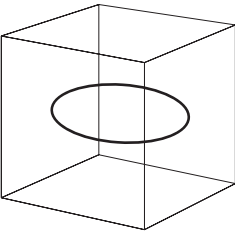




Initial view.



Structure view.



Camera path (thick line).

SCENE 1: ROTATION

In this scene, the viewer is inside of an enclosed cube which are also the data set's axes. The initial view appears to show that the cube is uniformly filled with data points, represented by shaded spheres. The size of the spheres is arbitrary to the data set and has been set at a scale that prevents the most distant data points from disappearing due to their distance from the camera.

In order to introduce the viewer to the 3d mouse, navigation in this scene is limited—the scene, not the camera, can only be rotated on an ellipse whose center point is also the center point of the enclosing cube. This means that directional illumination will alter shadows on each data point. Scenes 2 and 3 move the camera rather than the scene, adding to the feeling of presence and immersion.

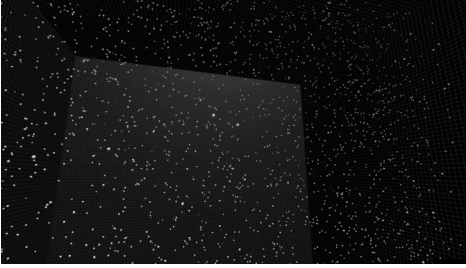
Once the viewer has rotated the scene 90 degrees (\pm approx. 10 degrees), the pattern of “banding” becomes apparent; data points are relatively constrained to the halfway point along one axis (in this case, the y -axis). Additionally, the head-mounted display provides depth cues through motion parallax as the viewer's head moves.



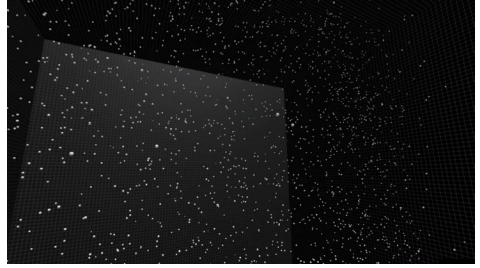
1 (initial view)



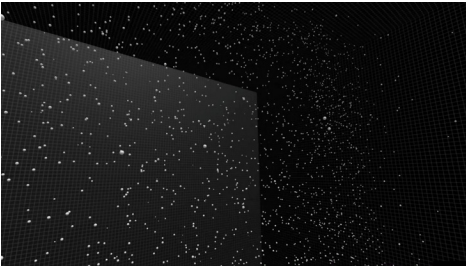
2



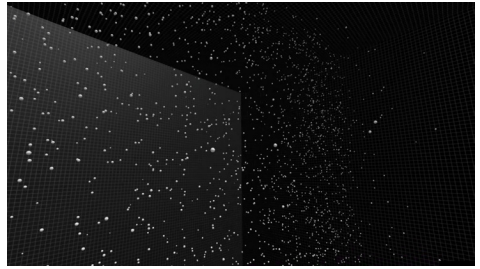
3



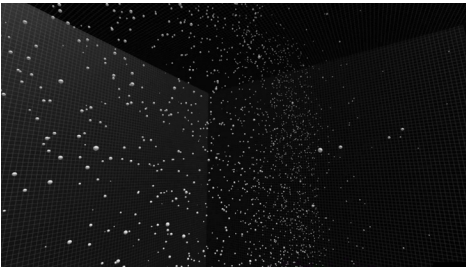
4



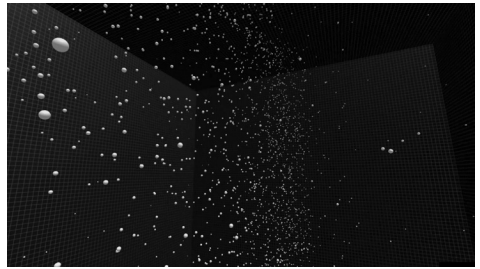
5



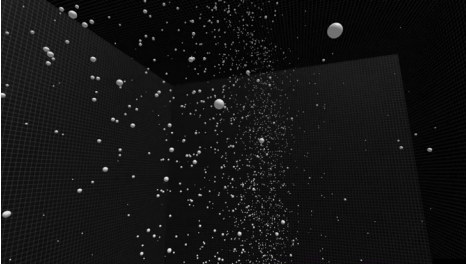
6



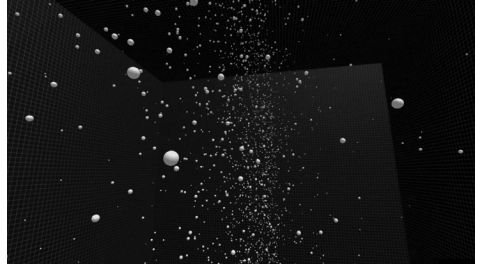
7



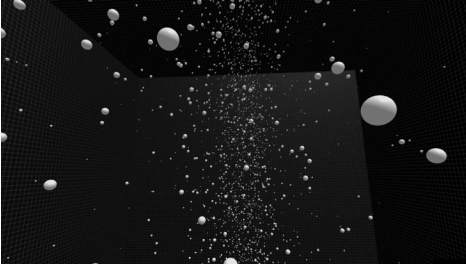
8



9



10



11 (structure view)

**SCENE 1: ROTATION
DATA POINT INITIALIZATION (SCRIPT RUNS MULTIPLE TIMES)**

```
using UnityEngine;
using System.Collections;

public class DataPointPrefab_Sphere_Random : MonoBehaviour {

    public GameObject prefab;
    public int quantity = 10;
    public float xMin = 0.0f;
    public float xMax = 1.0f;
    public float yMin = 0.0f;
    public float yMax = 1.0f;
    public float zMin = 0.0f;
    public float zMax = 1.0f;

    // Use this for initialization
    void Start () {
        for (int i = 0; i < quantity; i++) {
            float randX = Random.Range(xMin, xMax);
            float randY = Random.Range(yMin, yMax);
            float randZ = Random.Range(zMin, zMax);
            Vector3 pos = new Vector3(randX, randY, randZ);
            Instantiate(prefab, pos, Quaternion.identity);
        }
    }

    // Update is called once per frame
    void Update () {

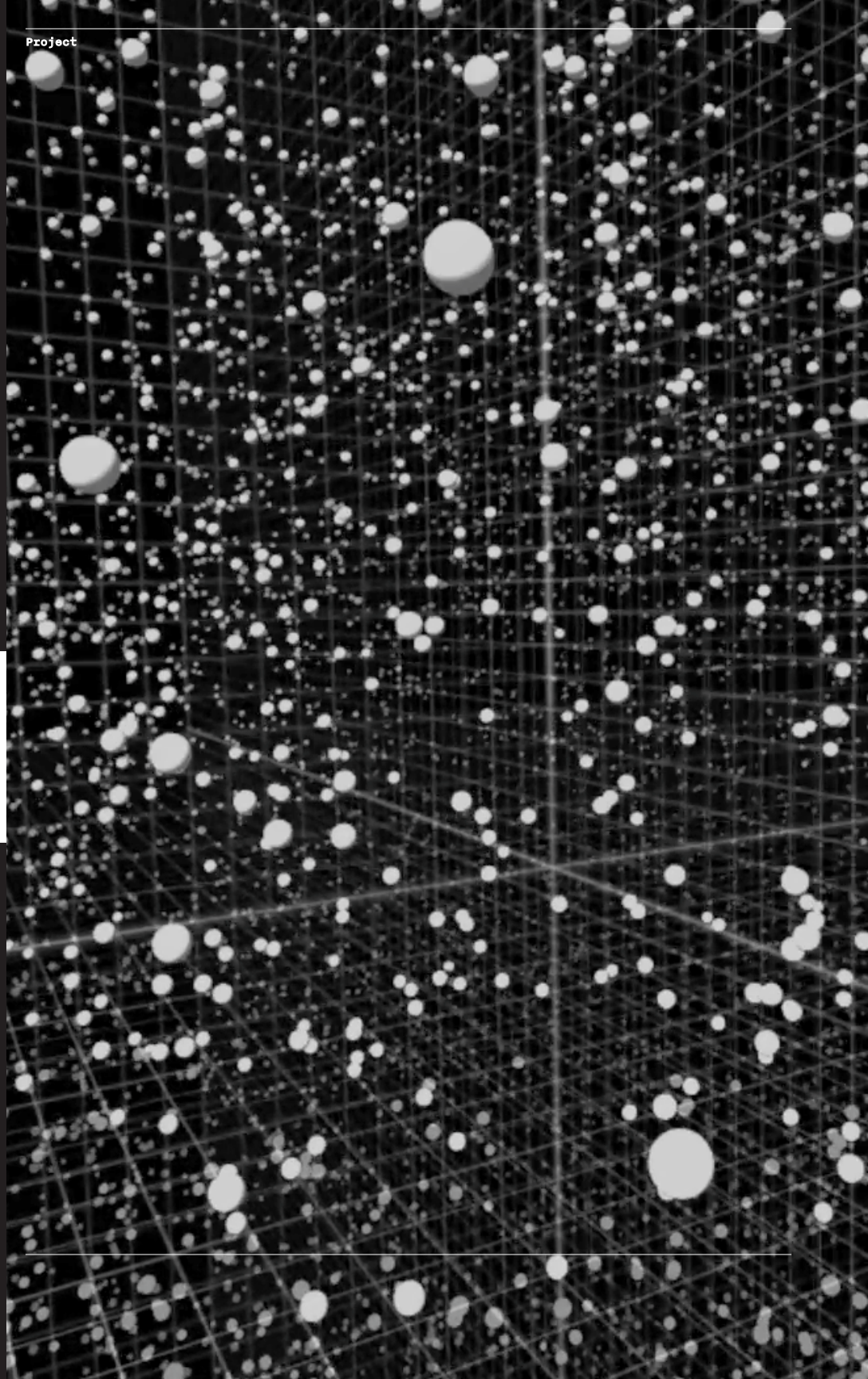
    }
}
```

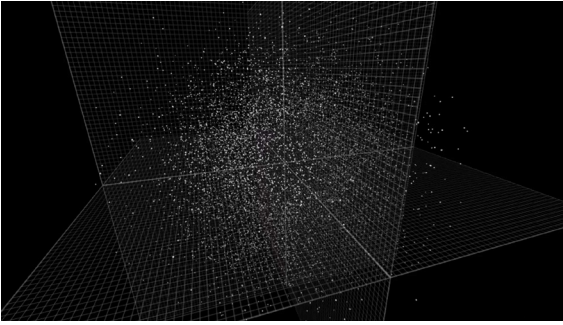
SCENE 1: ROTATION
CAMERA CONTROL LOCK (ATTACHED TO CAMERA OBJECT)

```
using UnityEngine;
using SpaceNavigatorDriver;

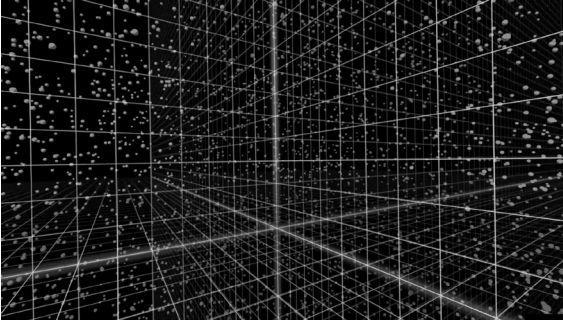
public class FlyAround_ThesisRotate : MonoBehaviour {

    public void Update () {
        transform.RotateAround(new Vector3(5f, 5f, 5f),
            Vector3.up, SpaceNavigator.Rotation.Yaw() *
            Mathf.Rad2Deg);
    }
}
```





Initial view.



Structure view.

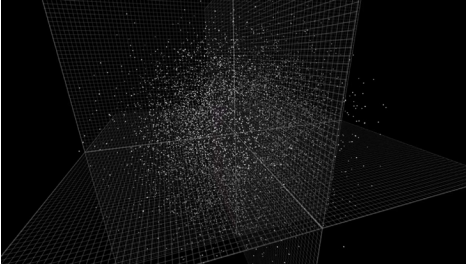
SCENE 2: SCALE

In this scene, the three axes (x-, y- and z-axis) are intersected rather than positioned to form an enclosure. This layout is needed for any data set with both positive and negative values, or if the origin (intersection of three axes) is needed to be positioned at a point other than [0, 0, 0].

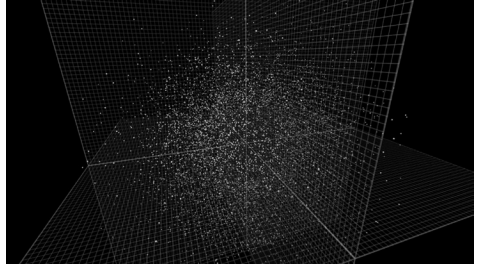
Control of the camera location is no longer constrained to a particular path, as in scene 1. Rather, the viewer is allowed to move freely in three dimensions. In order to prevent the viewer from turning the camera upside down—this can cause the viewer to feel lost—the camera is only allowed to rotate 90 degrees along the vertical (y) axis. Additionally, directional lighting assists in preserving the viewer's sense of the vertical. More simply put, which direction is up. The top of each sphere receives directional illumination from above the scene, similar to the sun shining from above.

In the initial view, the data set appears to uniformly occupy a sphere whose center is at the origin. Navigating the camera around the “outside” of the data set confirms this apparent structure. Once the viewer moves the camera near the origin, the hollow structure of the data set becomes apparent—there are no data points within a certain radial distance from the origin.

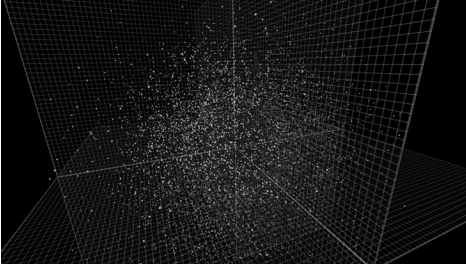
As in the first scene, the head-mounted display provides depth cues through motion parallax as the viewer's head moves. Additionally, if the viewer's head moves forward and backward, proximate data points that would otherwise obscure the view are omitted. In computer graphics, this is known as “clipping.”



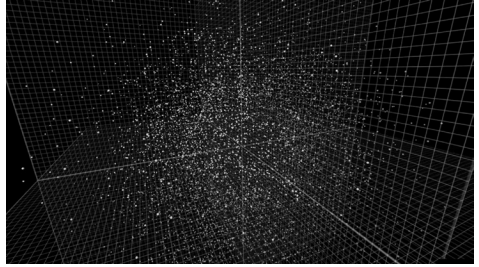
1 (initial view)



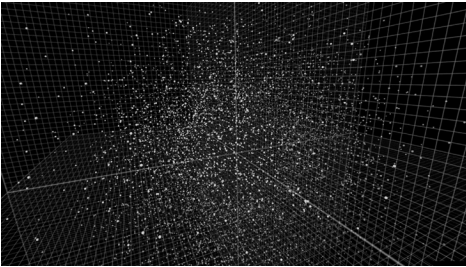
2



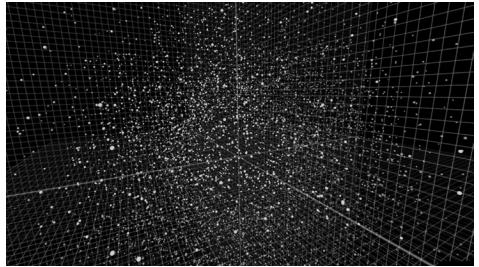
3



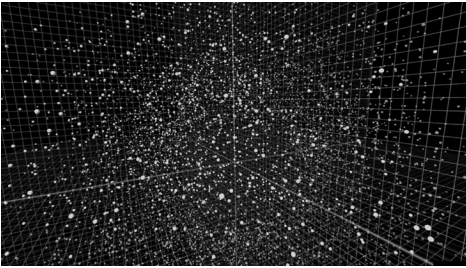
4



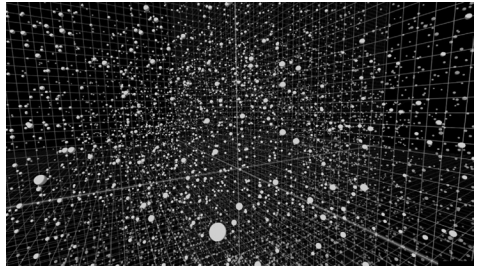
5



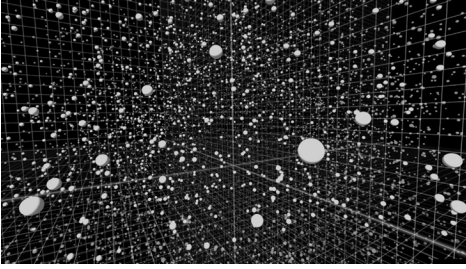
6



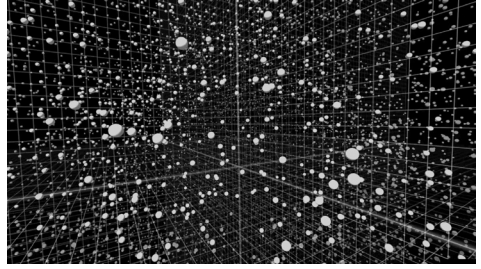
7



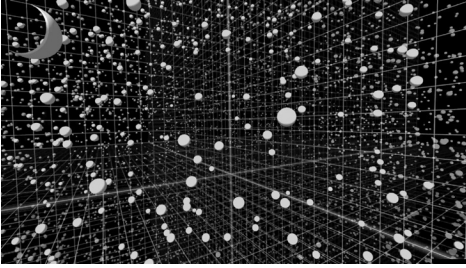
8



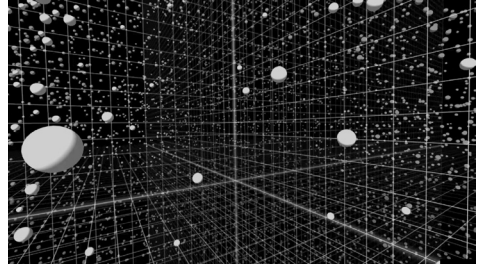
9



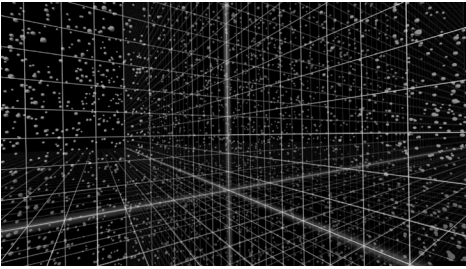
10



11 (clipping shown in top left)



12



13 (structure view)

SCENE 2: SCALE

DATA POINT INITIALIZATION (SCRIPT RUNS MULTIPLE TIMES)

```
using UnityEngine;
using System.Collections;

public class DataPointPrefab_Sphere_Random_Radial :
MonoBehaviour {

    public GameObject prefab;
    public int quantity = 10;
    public float minDistance = 2.5f;
    public float maxDistance = 3.5f;

    // Use this for initialization
    void Start () {
        for (int i = 0; i < quantity; i++) {
            Instantiate(prefab, Random.onUnitSphere * Random.
                Range(minDistance, maxDistance), Quaternion.
                identity);
        }
    }

    // Update is called once per frame
    void Update () {

    }
}
```

SCENE 2: SCALE

CAMERA CONTROL (ATTACHED TO CAMERA OBJECT)

```
using UnityEngine;
using SpaceNavigatorDriver;

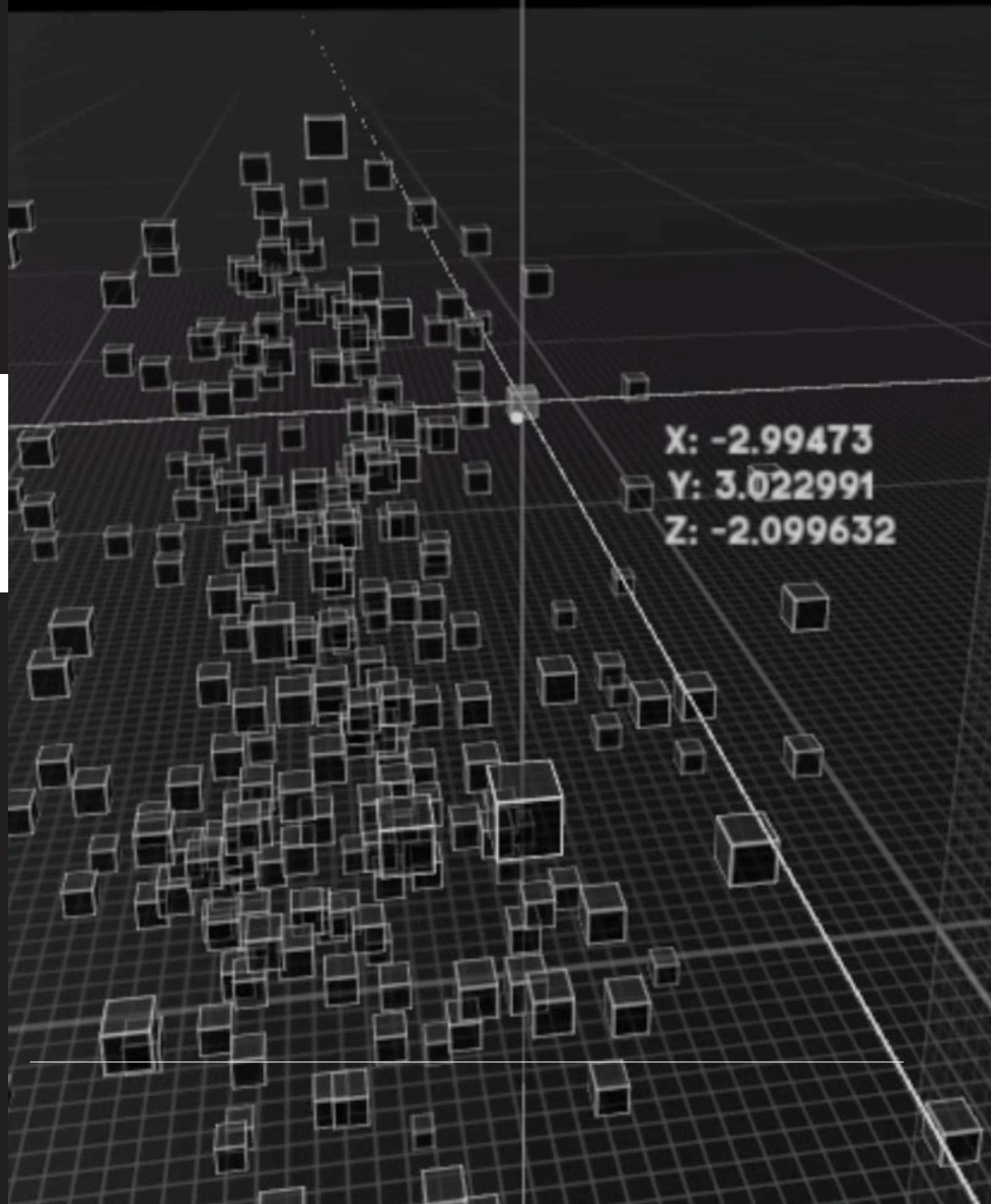
public class FlyAround : MonoBehaviour {
    public bool HorizonLock = true;

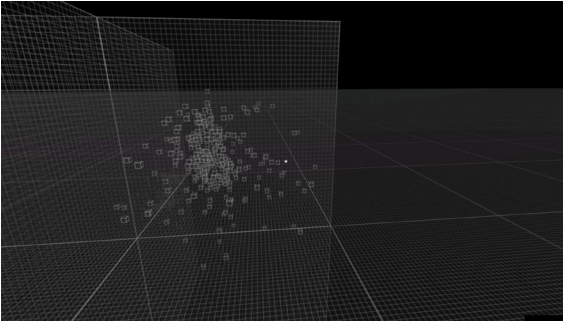
    public void Update () {
        transform.Translate(SpaceNavigator.Translation, Space.
            Self);

        if (HorizonLock) {
            // This method keeps the horizon horizontal

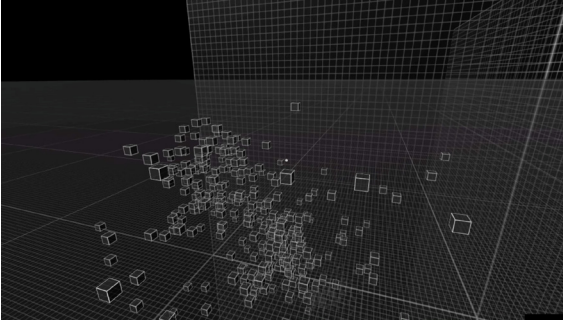
            // Perform azimuth in world coordinates.
            transform.Rotate(Vector3.up, SpaceNavigator.
                Rotation.Yaw() * Mathf.Rad2Deg, Space.World);

            // Perform pitch in local coordinates.
            transform.Rotate(Vector3.right, SpaceNavigator.
                Rotation.Pitch() * Mathf.Rad2Deg, Space.Self);
        } else {
            transform.Rotate(SpaceNavigator.Rotation.
                eulerAngles, Space.Self);
        }
    }
}
```





Initial view.



Structure view.

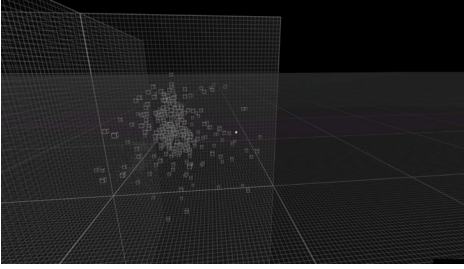
SCENE 3: ALIGNMENT

As in scene 2, the three axes (x-, y- and z-axis) are intersected. Building upon the theme of preserving the viewer's sense of the vertical—which direction is up—the x-axis has been extended into the horizon, serving as a ground plane. It is important to note that the viewer is still able to navigate the camera location below the ground plane which then becomes a "ceiling" plane. Still, both planes are quite easy to locate by looking up or down.

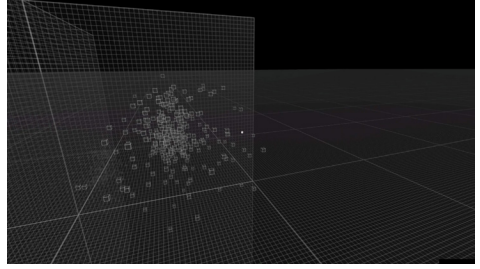
In the center of the viewer's view direction, a small reticle is displayed. This serves as yet another reference point, though the reticle is not completely static. Rather, the reticle will move slightly depending on the geometry of the object it hits—the reticle is a ray cast from the center of the field of view. When a data point is "hit" with the reticle, precise coordinates are displayed as well as perspective lines originating from the data points center.

Scene 3 also moves away from spheres as representations of individual data points. Instead, scene 3 uses cubes. As the cubes are also rendered in perspective, the edges provide trailing lines to the axis. This assist in estimating where along a particular axis a data point would lie. Direction lighting from above remains, though it is somewhat harder to discern from certain viewpoints.

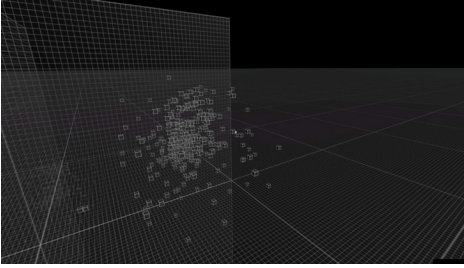
While these visual cues seem rather obvious, a surprising majority of 3d scatter plots ignore these visual cues as a provider of natural orientation. They are particularly important in VR as each cue makes extending viewing more comfortable.



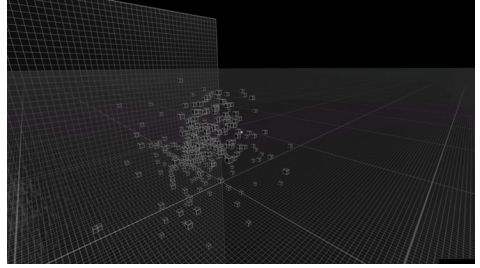
1 (initial view)



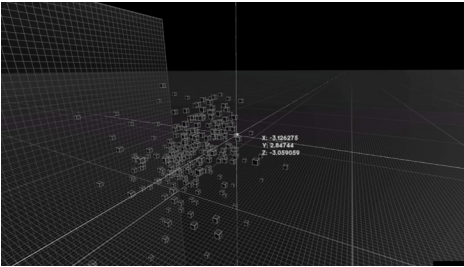
2



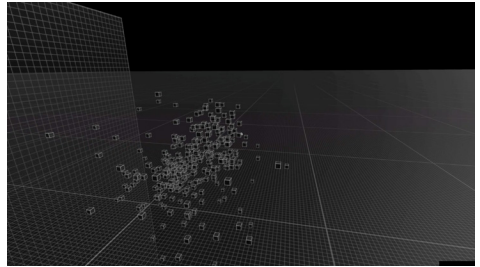
3



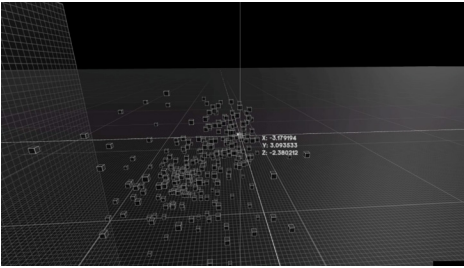
4



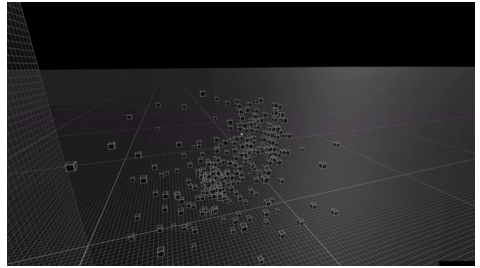
5



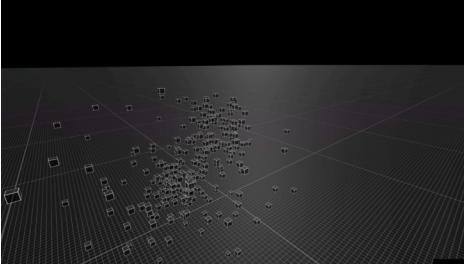
6



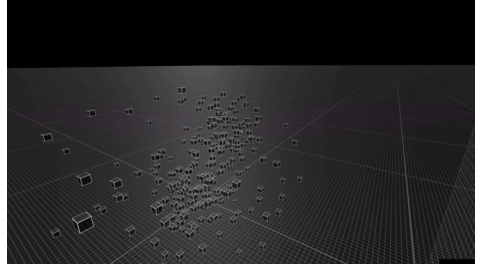
7



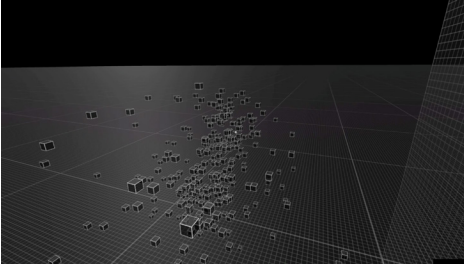
8



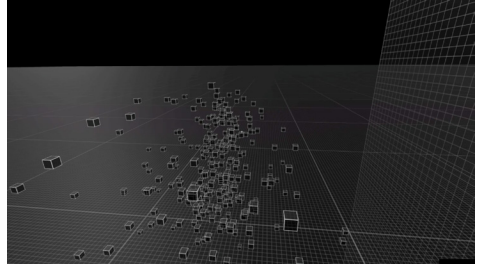
9



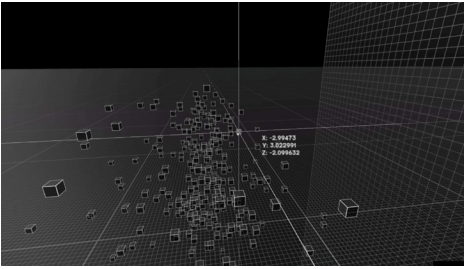
10



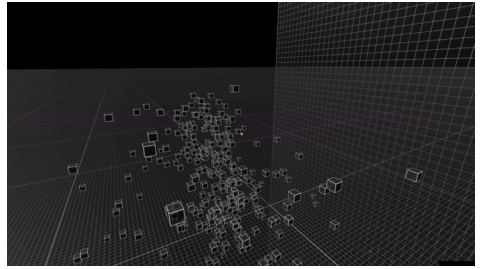
11



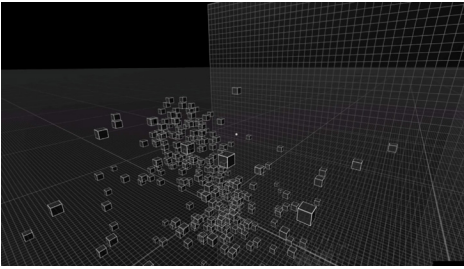
12



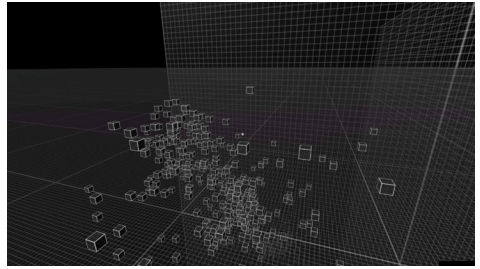
13



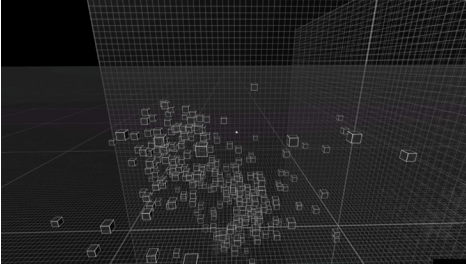
14



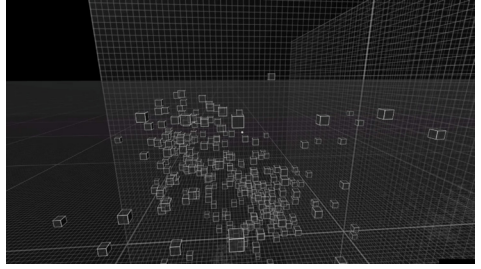
15



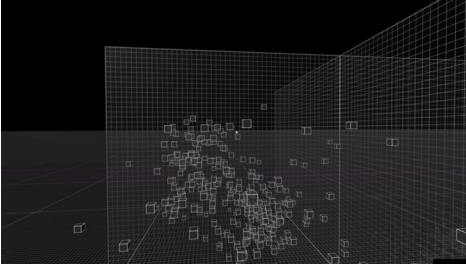
16



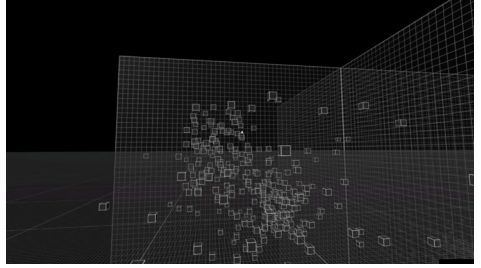
17



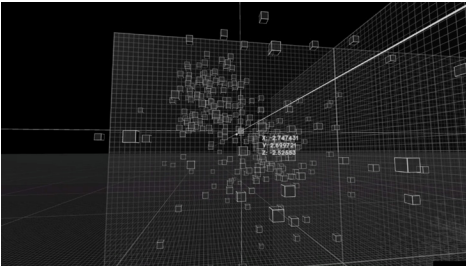
18



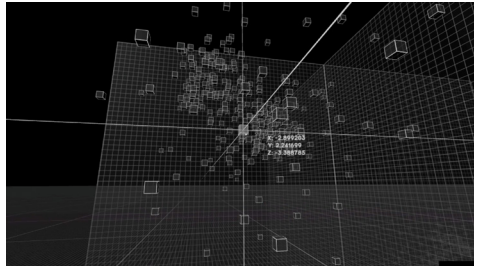
19



20



21



22 (structure view)

SCENE 3: ALIGNMENT
DATA POINT INITIALIZATION (SCRIPT RUNS MULTIPLE TIMES)

```
using UnityEngine;
using System.Collections;

public class DataPointPrefab_Cube_Random_Radial_Gaze :
    MonoBehaviour {

    public GameObject prefab;
    public int quantity = 10;
    public float minDistance = 2.5f;
    public float maxDistance = 3.5f;
    public float originX = 0f;
    public float originY = 0f;
    public float originZ = 0f;

    // Use this for initialization
    void Start () {
        for (int i = 0; i < quantity; i++) {
            Vector3 onSphere = Random.onUnitSphere * Random.
                Range(minDistance, maxDistance);
            Vector3 onSphereOrigin = onSphere + new
                Vector3(originX, originY, originZ);
            Instantiate(prefab, onSphereOrigin, Quaternion.
                identity);
        }
    }

    // Update is called once per frame
    void Update () {

    }
}
```

SCENE 3: ALIGNMENT
RETICLE RAYCAST COLLISION DETECT (ATTACHED TO CAMERA OBJECT)

```
using UnityEngine;
using System.Collections;

public class ShowDestinationPoint : MonoBehaviour {

    public GameObject destinationPoint;
    public Transform direction;

    // Use this for initialization
    void Start () {

    }

    // Update is called once per frame
    void Update () {
        RaycastHit hit;
        if (Physics.Raycast(direction.position, direction.
            forward, out hit)) {
            destinationPoint.transform.position = hit.
                point;
        }
    }
}
```

SCENE 3: ALIGNMENT
RETICLE RAYCAST COLLISION ACTION (ATTACHED TO CAMERA
OBJECT'S CenterEyeAnchor)

```
using UnityEngine;
using System.Collections;
using UnityEngine.UI;

public class TriggerByLooking : MonoBehaviour {

    public Transform direction;
    public float maxHitDistance;

    public GameObject selectLine;
    private GameObject xLine;
    private GameObject yLine;
    private GameObject zLine;

    private ChangeMaterial changeMaterial;
    private bool selecting;
    private GameObject lastObject;

    float x, y, z;
    public delegate void ShowPosition(float x, float y, float z);
    public static event ShowPosition showPosition;

    // Use this for initialization
    void Start () {
        selecting = false;
    }

    // Update is called once per frame
    void Update () {
        checkHit();
    }

    private void checkHit () {
        RaycastHit hit;
        if (Physics.Raycast(direction.position, direction.
            forward, out hit, maxHitDistance)) {
            if (hit.collider.gameObject.tag.Equals("Target")) {
                // Hit target object
                if (lastObject != null && !hit.collider.
                    gameObject.Equals(lastObject)) {
                    deselect();
                }
                if (!selecting) {
                    select(hit.collider.gameObject);
                }
            } else {
                // Hit non-target object
                if (selecting) {
                    deselect();
                }
            }
        }
    }
}
```

```
    }
  } else {
    // No object was hit
    if (selecting) {
      deselect();
    }
  }
}

private void select (GameObject hitObject) {
  selecting = true;
  changeMaterial = hitObject.GetComponent
    <ChangeMaterial>();
  changeMaterial.changeTo("DataPointPrefab_Cube_
    Material_Selected");
  lastObject = hitObject;
  x = hitObject.transform.position.x;
  y = hitObject.transform.position.y;
  z = hitObject.transform.position.z;
  if (showPosition != null) {
    showPosition(x, y, z);
  }
  Vector3 pos = new Vector3(x, y, z);
  xLine = (GameObject) Instantiate(selectLine, pos,
    Quaternion.identity);
  xLine.transform.Rotate(0, 90, 0);
  yLine = (GameObject) Instantiate(selectLine, pos,
    Quaternion.identity);
  yLine.transform.Rotate(90, 0, 0);
  zLine = (GameObject) Instantiate(selectLine, pos,
    Quaternion.identity);
}

private void deselect () {
  selecting = false;
  if (changeMaterial != null) {
    changeMaterial.changeTo("DataPointPrefab_Cube_
      Material");
  }
  UpdateInfoText.infoText.text = "";
  if (xLine) {
    Destroy(xLine);
  }
  if (yLine) {
    Destroy(yLine);
  }
  if (zLine) {
    Destroy(zLine);
  }
}
}
```

SCENE 3: ALIGNMENT
DATA POINT SELECT/DESELECT (ATTACHED TO DATA POINT
PREFAB OBJECT)

```
using UnityEngine;
using System.Collections;

public class ChangeMaterial : MonoBehaviour {

    public Material baseMaterial;
    public Material triggerMaterial;

    // Use this for initialization
    void Start () {

    }

    // Update is called once per frame
    void Update () {

    }

    public void changeTo(string mat) {
        if (mat.Equals("DataPointPrefab_Cube_Material")) {
            GetComponent<Renderer>().material = baseMaterial;
        }
        if (mat.Equals("DataPointPrefab_Cube_Material_
            Selected")) {
            GetComponent<Renderer>().material =
                triggerMaterial;
        }
    }
}
```



THE PERFECTION OF ART IS TO
CORCEAL ARTISTRY

UNIVERSITY OF BRISTOL
SCHOOL OF ART & DESIGN & DESIGN

THESIS

2018/19 + 2019

MAY 18 - JUNE 18

EXHIBITION

PAUL BURNHAM
DAN CLINE
CHLOE FRYLLA
JACOB HARRIS
HARRISON HALL GIBBERD
GEOFF JONES
YOUNG JAY
JANNA DE JONGHE
MELISSA MAGEE
MIRIAM MILES ROBERTSON
JAY SHARROD
SARAH SPENCER
EMMA TAYLOR HALL JONES
ELLEN TAYLOR JONES

EXHIBIT

While I have documented each of these three scenes as best as 2d printing allows, the core of this thesis is an interactive exhibit. For many visitors, this has been their first encounter with VR and for even more visitors, their first encounter with a 3d mouse. And while the 3d mouse is, for the most part, intuitive, there is still a slight learning curve. For myself, it took about an hour before I could navigate the camera exactly where I wanted it to go. Given that this project does not lead to brief, one-off encounters with a particular data set, the tradeoff is acceptable. Many video games have a higher learning curve, some upwards of 20 or 40 hours.

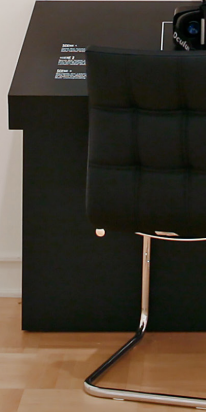
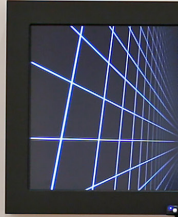
In the opening night of the exhibit, I stood near my project and offered minor pointers on how to control the camera. Rather surprising, the most overlooked and seemingly difficult interaction viewers had with the project was in changing scenes. The 3d mouse has two small buttons on either side of the base. Finding these buttons while wearing the HMD proved exceptionally difficult for some.

Perhaps the most surprising observation I made during the opening night was that there seemed to be two demographics that had no trouble at all: under 25 and over 60. For some reason, the 25-60 demographic, overall, experience difficulty not with the HMD but with the 3d mouse. This is understandable as it is intended primarily for people working with 3d modeling software.

NAVIGATING 3D SCATTER PLOTS IN IMMERSIVE VIRTUAL REALITY

OVERVIEW

OVERVIEW



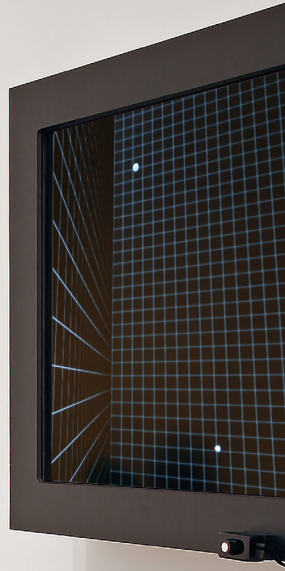


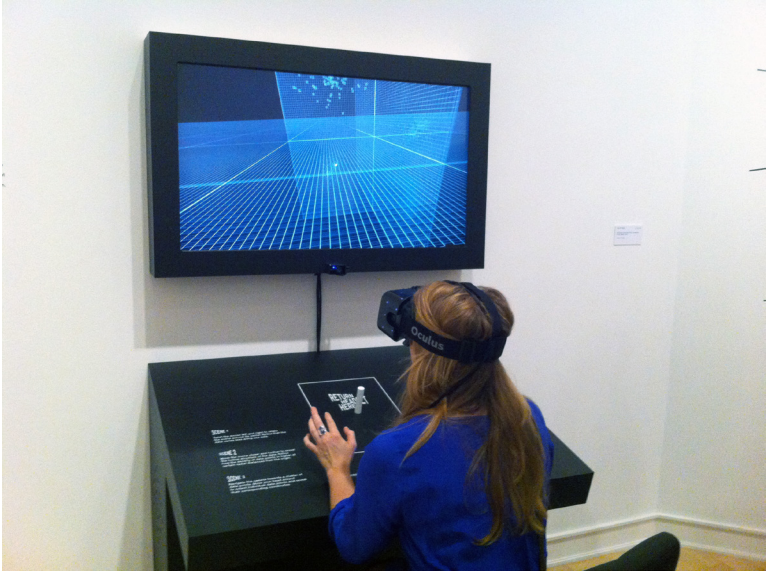
PROJECT

NAVIGATING 3D SCATTER PLOTS IN IMMERSIVE VIRTUAL REALITY

OVERVIEW

Virtual Reality (VR) is a technology that allows users to interact with a computer-generated environment. The technology is used in a variety of applications, including gaming, education, and training. In this project, we will explore how VR can be used to navigate 3D scatter plots in an immersive virtual reality environment. The project will focus on the design and development of a VR application that allows users to interact with 3D scatter plots in a virtual reality environment. The application will be designed to be user-friendly and easy to use, and will provide a comprehensive overview of the data represented by the scatter plots. The application will be developed using the Oculus VR platform, and will be tested on a variety of VR headsets. The project will also include a user manual and a series of tutorials to help users get started with the application. The project will be completed by the end of the semester, and will be presented to the class as a final project.







Conclusion

Even though Virtual Reality has been in development for decades, only now are industries dedicating sizable resources, both money and time, into producing compelling experiences for VR. Like many new technologies, entertainment has received the majority of attention. And while I was able to find a number of related projects, data visualization for VR is still in its infancy.

In particular, only a small number of projects explicitly concern themselves with visual principles that are relevant to developing for VR. From this project, there are a number of primary findings in this area, all of which can reduce the effects of motion sickness. Foremost is the inclusion of a ground plane. This static reference point is crucial for preventing disorientation and dislocation for the viewer. Also important are direction lighting from above, providing the viewer a regular visual indicator of which direction is up. Keeping points of interest, especially text to be read, in the center of the field of view is all but necessary. I anticipate new developments will occur in VR as organizations work more with the technology.

Given the growing demand for custom data visualizations, I believe VR will soon move away from the living room and into the office. Structures behind data sets are easily revealed in 3d display. This is not entirely surprising as our visual systems naturally operate in 3d space. Perhaps the greatest advantage to VR is motion

parallax. Additionally, the ability to “fly around” a data set can reveal surprising results to the data set’s structure. Patterns, clusters, banding, outliers and correlations are easy to spot within VR. Data visualization in virtual reality is truly and exploratory tool. I am especially excited to see how collaboration will come into play.

It remains to be seen if motion sickness can be completely removed from VR. We may find that acclimation is the only solution. Or we may find that there is no solution. If this is the case, I predict that mixed reality (MR) such as seen in Microsoft’s HoloLens will become the display of choice for immersive data visualization. This is an exciting area to be working in. I hope to continue this work well beyond my graduate education.



**Golden rule of
VR: The user
is in control of
the camera.**

Books: Human-Computer Interaction For Data Visualization

Buja, Andreas, John Alan McDonald, John Michalak, and Werner Stuetzle.

"Interactive Data Visualization Using Focusing and Linking." In *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on*, pp. 156-163. IEEE, 1991.

Murray, Scott. *Interactive Data Visualization for the Web*. O'Reilly Media, Inc., 2013.

—

Books: Data Visualization

Card, Stuart K., Jock D. Mackinlay, and Ben Shneiderman, eds. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1999.

Post, Frits H., Gregory Nielson, and Georges-Pierre Bonneau, eds. *Data Visualization: the State of the Art*. Vol. 1. Springer Science & Business Media, 2003.

Tufte, Edward R. "Envisioning Information." *Optometry & Vision Science* 68, no. 4 (1991): 322-324.

Tufte, Edward R., and P. R. Graves-Morris. *The Visual Display of Quantitative Information*. Vol. 2. Cheshire, CT: Graphics press, 1983.

Tufte, Edward R., and E. Weise Moeller. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Vol. 36. Cheshire, CT: Graphics Press, 1997.

Ware, Colin. *Visual Thinking: For Design*. Morgan Kaufmann, 2010.

—

Books: Human-Computer Interaction (HCI)

Soegaard, Mads, and Rikke Friis Dam. "The Encyclopedia of Human-Computer Interaction." *The Encyclopedia of Human-Computer Interaction*, 2012.

papers: human-computer interaction for data visualization

Drucker, Steven M., Danyel Fisher, Ramik Sadana, and Jessica Herron. "TouchViz: a Case Study Comparing Two Interfaces for Data Analytics on Tablets." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2301-2310. ACM, 2013.

Espinosa, Octavio Juarez, Chris Hendrickson, and J. H. Garrett. "Domain Analysis: a Technique to Design a User-centered Visualization Framework." In *Information Visualization, 1999. (Info Vis '99) Proceedings. 1999 IEEE Symposium on*, pp. 44-52. IEEE, 1999.

Few, Stephen. "Data visualization for human perception." *The Encyclopedia of Human-Computer Interaction*, 2nd Ed., 2013.

Few, Stephen. *Information Dashboard Design*. O'Reilly, 2006.

Few, Stephen. *Now You See it: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.

Few, Stephen. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Vol. 1, no. 1. Oakland, CA: Analytics Press, 2004.

Heer, Jeffrey, and Ben Shneiderman. "Interactive Dynamics for Visual Analysis." *Queue* 10, no. 2 (2012): 30.

Heer, Jeffrey, Maneesh Agrawala, and Wesley Willett. "Generalized Selection Via Interactive Query Relaxation." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 959-968. ACM, 2008.

Heer, Jeffrey, Stuart K. Card, and James A. Landay. "Prefuse: a Toolkit for Interactive Information Visualization." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 421-430. ACM, 2005.

Hochheiser, Harry, and Ben Shneiderman. "Visual Queries for Finding Patterns in Time Series Data." *University of Maryland, Computer Science Dept. Tech Report, CS-TR-4365*, 2002.

Khan, Muzammil, and Sarwar Shah Khan. "Data and Information Visualization Methods, and Interactive Mechanisms: a Survey." *International Journal of Computer Applications* 34, no. 1 (2011): 1-14. Langton, John T., Elizabeth A. Gifford, and Timothy J. Hickey. "Visualization and Interactive Exploration of Large, Multidimensional Data Sets." In *Applications of Computational Intelligence in Biology*, pp. 231-255. Springer Berlin Heidelberg, 2008.

Robertson, George G., Stuart K. Card, and Jack D. Mackinlay. "Information Visualization Using 3D Interactive Animation." *Communications of the ACM* 36, no. 4, 1993: 57-71.

Rzeszotarski, Jeffrey M., and Aniket Kittur. "Kinetica: Naturalistic Multi-touch Data Visualization." In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pp. 897-906. ACM, 2014.

Sadana, Ramik, and John Stasko. "Interacting with Data Visualizations on Tablets and Phones: Developing Effective Touch-based Gestures and Operations."

Satyanarayan, Arvind, Kanit Wongsuphasawat, and Jeffrey Heer. "Declarative interaction design for data visualization." In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, pp. 669-678. ACM, 2014.

Seo, Jinwook, and Ben Shneiderman. "A Rank-by-feature Framework for Interactive Exploration of Multidimensional Data." *Information Visualization* 4, no. 2, 2005: 96-113.

Yi, Ji Soo, Youn ah Kang, John T. Stasko, and Julie A. Jacko. "Toward a Deeper Understanding of the Role of Interaction in Information Visualization." *Visualization and Computer Graphics*, IEEE Transactions on 13, no. 6, 2007: 1224-1231.

—

Papers: Data Visualization (General)

Card, Stuart K., and Jock Mackinlay. "The Structure of the Information Visualization Design Space." In *Information Visualization*, 1997. Proceedings, IEEE Symposium on, pp. 92-99. IEEE, 1997.

Cleveland, William S., and Robert McGill. "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." *Journal of the American Statistical Association* 79, no. 387, 1984: 531-554.

Healey, Christopher G. "Perception in Visualization." <http://www.csc.ncsu.edu/faculty/healey/PP/index.html>, 2007.

Rodrigues, Jose F., Agma JM Traina, and Maria Cristina Ferreira de Oliveira. "Reviewing Data Visualization: an Analytical Taxonomical Study." In *Information Visualization*, 2006. IV 2006. Tenth International Conference on, pp. 713-720. IEEE, 2006.

Shneiderman, Ben. "The Eyes Have It: A Task By Data Type Taxonomy for Information Visualizations." In *Visual Languages*, 1996. Proceedings., IEEE Symposium on, pp. 336-343. IEEE, 1996.

Tukey, John W. "Exploratory Data Analysis." Reading, Ma 231 (1977): 32.

—

Papers: Data Visualization (Specific Studies)

Acevedo, Daniel, Cullen D. Jackson, Fritz Drury, and David H. Laidlaw. "Using Visual Design Experts in Critique-based Evaluation of 2D Vector Visualization Methods." *Visualization and Computer Graphics*, IEEE Transactions on 14, no. 4 (2008): 877-884.

Donalek, Ciro, S. George Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh et al. "Immersive and collaborative data visualization using virtual reality platforms." In *Big Data (Big Data)*, 2014 IEEE International Conference on, pp. 609-614. IEEE, 2014.

Hahn, Horst K., Bernhard Preim, Dirk Selle, and H-O. Peitgen. "Visualization and Interaction Techniques for the Exploration of Vascular Structures." In *Visualization*, 2001. VIS'01. Proceedings, pp. 395-578. IEEE, 2001.

Heer, Jeffrey, and Michael Bostock. "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 203-212. ACM, 2010.

Jarosch, Rafael, Grund, Sebastian, and Sprott, Sascha. "Interactive Exploration of Immersive Illuminated 3D Scatterplots."

Langton, John T., Astrid A. Prinz, and Timothy J. Hickey. "NeuroVis: Combining Dimensional Stacking and Pixelization to Visually Explore, Analyze, and Mine Multidimensional Multivariate Data." In *Electronic Imaging 2007*, pp. 64950H-64950H. International Society for Optics and Photonics, 2007.

Nagel, Henrik R., Erik Granum, and Peter Musaeus. "Methods for visual mining of data in virtual reality." In Proceedings of the International Workshop on Visual Data Mining, pp. 13-27. 2001.

—

Papers: Human-Computer Interaction (HCI)

Robertson, George, Stuart K. Card, and Jock D. Mackinlay. "The Cognitive Coprocessor Architecture for Interactive User Interfaces." In Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology, pp. 10-18. ACM, 1989.

—

Books/Papers: Perception

Camurri, Antonio, and Marc Leman. *Gestalt-based Composition and Performance in Multimodal Environments*. Springer Berlin Heidelberg, 1997.

Durand, Frédo, Maneesh Agrawala, Bruce Gooch, Victoria Interrante, Victor Ostromoukhov, and Denis Zorin. "Perceptual and Artistic Principles for Effective Computer Depiction." *Siggraph 2002 Course #13 Notes* (2002).

Gibson, James J. *The ecological approach to visual perception: classic edition*. Psychology Press, 2014.

References

Koffka, Kurt. Principles of Gestalt Psychology. Routledge, 2013.

Stevens, Staley S. "The Psychophysics of Sensory Function." American Scientist (1960): 226-253.

—

Other

UW Interactive Data Lab (Jeff Heer)

Cecilia Aragon (UW HCDE, Associate Professor, HCI, CSCW, Visual Analytics, Director | Human-Centered Data Science Lab)

<http://bl.ocks.org>. (mbostock's (Mike Bostock) blocks)

<http://d3.js>

<http://unity3d.com>

<http://datavisualization.ch> (data vis for the web, tools)

<http://flowingdata.com> (Nathan Yau, visualizing data author)

<http://vis.berkeley.edu> (Visualization Lab, UC, Berkeley)

Visualization and Behavior Group (IBM Research)

WatsonAnalytics (IBM)

Many Eyes (IBM)

<http://mkweb.bcgsc.ca> (Martin Krzywinski)

Color Brewer 2.0 (colorblind-safe color palettes)

<http://vwwv.org> (a multipurpose toolkit)

<http://processing.org>

<http://puredata.info>

Max/MSP

Fernanda Viégas & Martin Wattenberg

David McCandless (Data Journalist)

AT&T Labs Research

SIDL (Spatial Information Design Lab, Graduate School of Architecture, Planning and Preservation at Columbia University)

JoAnn Kuchera-Morin (CREATE, The Center for Research in Electronic Art Technology, UC, Santa Barbara)

<http://ieevis.org> (CS, CG, DV)

<http://do.minikus> (Dominikus Baur, Data Vis & Mobile Interaction Designer)

Ig Nobel Prize, Improbable Research

Virtual reality reveals spatially complex structures behind 3d data in an easily explorable way. The difficulty is in navigating this space while wearing a head-mounted display (HMD). One viable option for view control is a combination of a 3d mouse for camera location and head tracking of the HMD for the control of view direction. The scenes shown here display semi-randomized 3d scatter plots, each of which emphasizes a type of navigation that leads the viewer to discover structure behind the data. The three types of navigation are rotation, scale and alignment.