FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Immersive 2D Scatterplots and Parallel Coordinates for Information Visualization in a Virtual Reality Environment

Paulo Jorge Palhau Moutinho

Mestrado em Engenharia Informática e Computação

Supervisor: Alexandre Valle de Carvalho

Supervisor: Daniel Mendes

Supervisor: Rui Rodrigues

July 25, 2022

Immersive 2D Scatterplots and Parallel Coordinates for Information Visualization in a Virtual Reality Environment

Paulo Jorge Palhau Moutinho

Mestrado em Engenharia Informática e Computação

July 25, 2022

Abstract

The analysis of vast volumes of data can be an unpleasant and difficult endeavor. Sometimes the causal relationships between events are not easy to find by looking at individual pieces of information. The way information is presented is critical to its accurate and easier understanding. By encompassing more pieces of information from multiple sources onto a single coherent visual representation the user may be able to identify causal relations among other extracted knowledge.

Information Visualization is the study of techniques for interactive and visual representation of complex data in a manner that best highlight the data structure and respective causal relations to ease in the process of its understanding. Over the years, it has been coupled with Virtual Reality, which is considered an alternative medium for both 2D and 3D representations by allowing the user to engage with the information in a natural and immersive way.

A field that deals with large data volumes and benefits from proper visualization tools is the financial domain. Geopolitical, social and economic events have an impact on the performance of financial assets. For a private investor, understanding how these events are related is a complex and time-consuming process. An immersive visualization of both asset performance chart and relevant financial indicators might help discern what's influencing value change.

We developed a solution that encompasses techniques from the field of Information Visualization with the field of Immersive Environments, considering the specific context of financial data. Within this study, a prototype was developed and tested, to ascertain if the proposed solution helps closing the gap between the users and the underlying patterns existing in the data. The results show the proposed solution is more intuitive, appealing and easy to use that a Desktop setup in tasks concerning data analysis, although it can be slower at times.

Keywords: Information Visualization, Virtual Reality, 2D Scatterplots, Parallel Coordinates Plots, Immersive Environments, Data Analysis

Resumo

A análise de vastos volumes de dados pode ser uma tarefa desagradável e difícil. Por vezes, as relações causais entre eventos não são fáceis de encontrar através da análise de fontes de informação individuais. A forma como a informação é apresentada é fundamental para a sua compreensão precisa e mais fácil. Ao englobar mais informação de múltiplas fontes numa única representação visual coerente, o utilizador pode ser capaz de identificar as relações causais entre outros conhecimentos extraídos.

Visualização da Informação é a área de conhecimento que estuda as técnicas de representação interactiva e visual de dados complexos de uma forma que melhor realce a estrutura de dados e respectivas relações causais para facilitar a sua compreensão. Ao longo dos anos, tem sido associada à Realidade Virtual, que é um meio alternativo considerado tanto para representações 2D como 3D, permitindo ao utilizador envolver-se com a informação de uma forma natural e imersiva.

Uma área que trabalha com vastos volumes de dados e beneficía de ferramentas adequadas de visualização é o domínio financeiro. Os eventos geopolíticos, sociais e económicos têm um impacto sobre o desempenho de ativos financeiros. Para um investidor privado, compreender como estes eventos estão relacionados é um processo complexo e demorado. Visualizar o gráfico do desempenho de ativos juntamente com indicadores financeiros relevantes pode ajudar a discernir o que está a influenciar a mudança de valor.

Deste modo, desenvolvemos uma solução que engloba técnicas do ramo da Visualização da Informação com o domínio de Ambientes Imersivos, considerando o contexto específico dos dados financeiros. Dentro desta solução, foi desenvolvido e testado um protótipo, para verificar se a solução proposta ajuda a aproximar os utilizadores dos padrões presentes nos dados. Os resultados mostram que a solução proposta é mais intuitiva, apelativa e fácil de usar em tarefas relativas à análise de dados, apesar de mais lenta por vezes.

Keywords: Visualização de Informação, Realidade Virtual, Gráficos 2D, Gráficos de Coordenadas Paralelas, Ambientes Imersivos, Análise de Dados

Acknowledgments

I must sincerely thank the support, encouragement, and incentive given to me during the course of this dissertation. I would like to thank the level of dedication shown and advice provided by my supervisors, Prof. Alexandre Valle de Carvalho, Prof. Daniel Mendes, and Prof. Rui Rodrigues.

I would like to show my gratitude for the excellent atmosphere I was provided during the course of this dissertation at the GIG laboratory. I also would like to thank FEUP for the opportunity and necessary means to develop this work.

To my friends, I would like to express my sincere appreciation for all the patience, advice, distraction, and assistance given. With you, I share the success of this accomplishment.

I would also like to thank everyone who gladly participated in the user evaluation sessions.

Also, a special thanks to my parents and sister for the constant support and encouragement during the course of another chapter of my life.

Paulo Jorge Palhau Moutinho

“The message meaning that the divine gift does not come from a higher power but from our own minds.“

Anthony Hopkins

“There’s room for a lot in this world and if you have a whole planet of carpenters, then nothing ever gets done. You need a bit of everything and we’re fine.“

Caduceus Clay

“Life is a game and you are making it.“

Miguel Pinto

Contents

List of Figures

List of Tables

Abbreviations

CAVE	Cave Automatic Virtual Environment
Infoviz	Information Visualization
HMD	Head-mounted Display
PCP	Parallel Coordinates Plot
STC	Space-time Cube
SUS	System Usability Scale
VR	Virtual Reality
VRE	Virtual Reality Environment
WTI	Western Texas Intermediate

Chapter 1

Introduction

With the recent data boom, there has been an effort to organize and categorize the seemingly unlimited amount of data available. People nowadays have access to a wide range of data, from high-level and broad statistics to up-to-the-minute news. However, understanding how these relate to real-life events is a complex and time-consuming process.

Immersive Analytics is a research initiative, whose goal is aimed at bridging the gap between people and information they analyze to draw conclusions [40]. It encompasses the field of Information Visualization, commonly referred to as InfoViz, which examines novel ways for people to view and interact with abstract data, and Virtual Reality (VR), which allows for an immersive experience while interacting with the information, through its visual representation.

1.1 Context and Motivation

Geopolitical, financial, and social events influence our lives every day. The general public has access to information regarding these events in a plethora of ways, however, due to its amount, different levels of detail, outward indifference, and the way it is visualized, finding the causal relationships might be more than what the casual user can handle. By structuring these occurrences and grouping them in a proper visualizing medium, relationships and patterns become more easier to find, which potentially can help draw conclusions and support the decision-making process.

The concept of Information Visualization has long piqued people's interest [22], [17]: creating interactive and visual representations of complex data to reinforce human cognition by allowing viewers to learn about the data's structural organization and causal relations.

In 1987, it was strongly linked to Scientific Computing and described as providing " a method for seeing the unseen. It enriches the process of scientific discovery and fosters profound and expected insights. In many fields, it is already revolutionizing the way scientists do science" [41]. At the same time, Edward Tuftle, a seminal author of the field, pioneered the discussion on several topics of data visualization that are still considered relevant to this day [50]. Furthermore, the benefits of Information Visualization have been researched in a variety of fields, including finances, business, health, marketing, among other domains that deal with large, complex and abstract data.

The way people interact with information can be as important as the devised visual representation itself. Mostly since the 2000s, new types of interaction technique has been explored, post-WIMP (post windows, icons, mouse, and pointers), which attempts to push beyond the limitations of 2D interfaces. In light of this, InfoVis has been combined with Virtual Reality to research new ways of representing and interacting [52], [25], [48].

Virtual Reality allows the creation of immersive environments, which may be a better medium for 3D display by letting the user see the information in a broader way as well as a 2D display by virtue of being able to present spacious and interactive 2D surfaces [49]. Current VR systems are becoming more capable as hardware also becomes more powerful. They include HMDs (Head-mounted Displays, available since the 1960s), tracking systems, and handheld input devices that are both reliable and simple to set up. Furthermore, today's computers have significantly improved graphics capability, reducing performance difficulties.

The use of proper visualization techniques allow the user to better understand the complex data structures, enable potential for the discovery of causal patterns and may lead to a better understanding of the information and possibly the drawing of conclusions that would be unnoticed otherwise.

1.2 Problem

Multiple sources of data should be analyzed before the taking of any decision in order to reach solid conclusions and better justify our choices. The manner in which the information is presented and interacted with is crucial to its easier and more accurate understanding. However, its a complex and time-consuming process going through various and distinct pieces of information, specially if their representations are independent from each other, without showing how they are connected. A good example of this happens regularly in the financial domain.

Finance investors can research and formulate their opinions on a very wide set of heterogeneous sources such as finance websites, newspapers, economic indexes and rates, etc. However, the private investor might not have the same amount of analysis resources or the required time to go through each data source. Sometimes, the current state of the financial markets can be hard to understand since its quite a complex system full of input variables and impacted by the continuous flow of economic, financial and geopolitical events across the globe. For instance, given the recent past, what is the impact in the stock market of the FED governor declaring rates will be increased shortly? What is the impact on gas commodities when there is a potential Russian invasion of Ukraine?

By visualizing diversified pieces of information separately, the process of finding their relations will require more time with the possibility of reducing the extracted knowledge and achieving rushed conclusions.

1.3 Hypothesis

Using Information Visualization techniques, we propose a visual narrative that encompasses multiple sources of data (charts, notes, streams, opinions, facts) with the goal of understanding the possible inter-relationships through the data. The multiple pieces of information should be represented coherently, providing as immersive experience when the user interacts with the data.

Considering the financial domain, we expect the user to better perceive the behavior of a market without having to "drill down" the numerous sources available (mostly text and 2D charts), which would be impractical, if not impossible, due to the number of updates, volume, and heterogeneity.

1.4 Objectives

Our goal is to research and propose a solution towards the visual display of quantitative information from a specific data domain, considering an immersive environment as part of the interface, and using adequate visualization techniques that reinforce the structural relations within the data. We design these representations using a set of Virtual Reality interaction techniques with the goal to provide the most natural and immersive experience. Overall the goals are:

a) to design one or more visualization techniques that best suit the domain of data chosen as well as the highlight of the causal relationships present.

b) to design an interaction model, targeting the visual representations, based on locomotion and selection techniques that can provide a good immersive and intuitive experience.

c) to prototype the proposed solution and evaluate it by conducting tests with users mostly through lab experiment and questionnaires.

1.5 Document Structure

The document is composed of multiple sections: the second chapter presents the surveyed state of the art. In chapter three, Proposed Solution, we describe the base concepts of our solution. We then move on to our proposed technique, architecture, hardware as well as describing several decisions behind the prototype implementation, in chapter four. Next, chapter five presents the methodology used for experimentation and the metrics we used to measure the performed experiments. In this chapter, we talk about the dissertation's expected execution plan and we will present the results and consequent discussion.

The last chapter presents the main conclusions about our work, addresses the original problem and goals achieved. It also presents threads regarding future work.

Chapter 2

State of the Art

To properly contextualize our work, we will discuss existing literature regarding Visualization Techniques, with a focus on those present in VR environments, Immersion and how it is accomplished. Finally, some important aspects are also presented in the area of Finance visualization.

2.1 Visualization on Virtual Reality Environments

Over time, there has been an increasing effort to determine which immersive visualization techniques perform best with the various types of data [37], [24]. We will introduce visualization approaches developed in VR environments, especially focusing on those we consider better for displaying high-dimensional data, as well as visualization tasks that are used to ascertain how well users perform during the visualizations.

2.1.1 Multidimensional Visualizations

With the introduction of VR environments, some traditional as well as novel 3D techniques are again being explored due to the lack of limitations this medium has in comparison to screen-based ones. Thus, we showcase some multidimensional display techniques of information, such as scatterplots, parallel coordinates, geographic oriented visualizations, glyphs, symbols and icons, which are being explored in immersive environments.

Scatterplots

Also known as scatterplot matrices, are mathematical diagrams that represent information through the display of points positioned in three orthogonal axes [34].

In a comparison study, Filho *et al.* [26] examined multidimensional reduction data in desktop-based 2D and 3D and HMD-based 3D representations. The goal was to see how users would respond to each one and demonstrate the benefits of 3D visualizations in immersive environments.

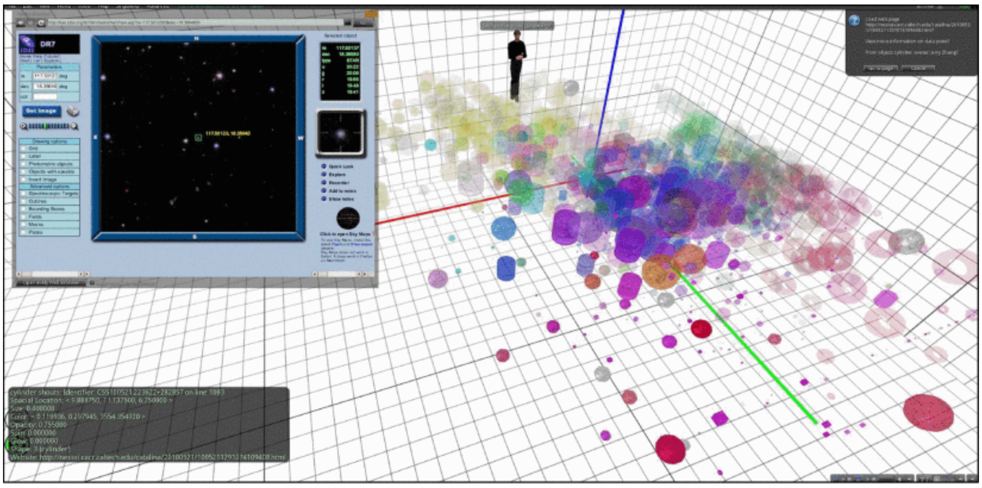


Figure 2.1: High-dimensional Scatterplot data visualization in VRE [25]. Color, size, transparency, shape used for additional dimensionality

Users performed much better in VR in distance perception and outliers identification tasks, which can be explained by the visible and interaction constraints of desktop-based devices. However, when it comes to user experience, users found the desktop-based setup to be more intuitive and faster, while the HMD-based setup to be more engaging. The authors justify these ratings with a possible unfamiliarity with these types of new environments.

Taking advantages of multiple configurations regarding visual quality, Donalek *et al.* [25] experimented with the colors, shapes, sizes, transparencies, and textures of each point to generate eight-dimensional scatterplots (see Figure 2.1). A prototype was created that allowed several users to visualize the data and collaborate at the same time. The way data was represented through the eight dimensions was helpful in terms of user comprehension, and the ability to dynamically adjust visual aspects during the visualization considerably increased user engagement and immersion.

Arms *et al.* [15] investigated the performance of structural tasks, one of which being cluster identification in virtual environments and desktop-based devices. The clusters were identified using colors and glyphs, and the results showed better performance in VR, which the authors justified as these environments provide better 3D visualization by immersing the users in the data.

There are few frameworks that allow quick prototyping of Scatterplots in VRE. ImAxes [23], an authoring tool, allows the creation of visualization techniques dependent on axes by requiring the use of simple programming grammar and also providing templates, which can be customized and turned into working prototypes (see Figure 2.2).

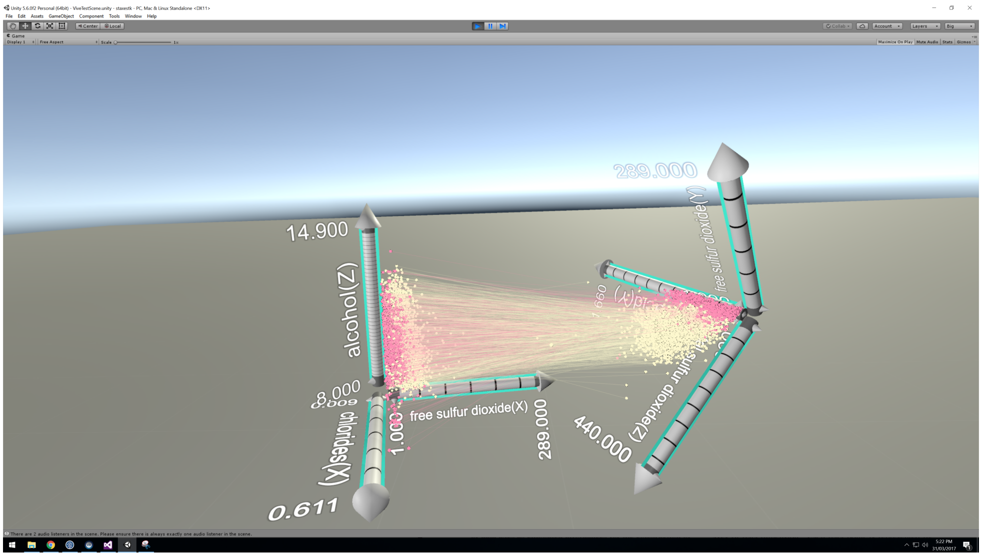


Figure 2.2: Tree-linked representation of Scatterplots produced by the Infoviz framework ImAxes [23]

In a similar way, DXR [45], an authoring tool, facilitates the development of high-dimensional visualization techniques, including Scatterplots, and offers a degree of personalization in several elements present in the representation (see example in Figure 2.3).

Parallel Coordinates

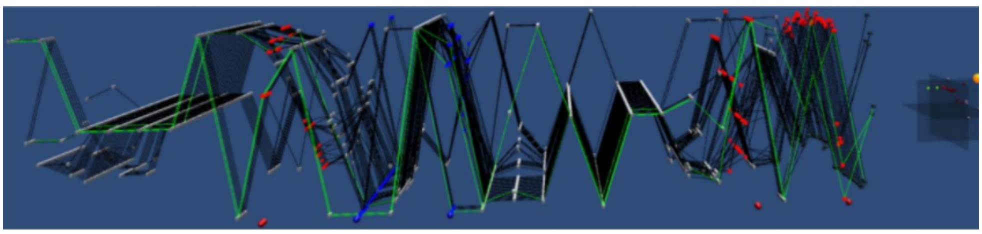


Figure 2.3: High-dimensional Parallel Coordinates Plot visualization in VRE. Colours are used to highlight and more easily track different points through all dimensions [48]

Parallel coordinates plots, or PCPs, are a common way of viewing and further analyze large and complex datasets. It represents the data through points placed on parallel axes (see Figure 2.3). This visualization technique overcomes scatterplots when intending to trace a singular data point in several dimensions. In Scatterplots, that is only possible in adjacent planes.

Billow and Cottam [19] developed Geovisor, a tool that allows users to create abstract 3D visualization approaches, and used it to analyze user performance evaluation criteria in a variety of representations, including PCPs, with the purpose of detecting correlated data.

Tadeja *et al.* [48] examined the potential benefits of using PCPs to aid with engineering design decision-making. They used a dataset that had previously been examined with 2D PCPs to see if using 3 dimensions with the help of virtual reality would result in improved data analysis outcomes. Users with domain expertise and experience visualizing and interacting with PCPs were asked to identify patterns in the dataset. There was some discrepancy in performance among the four users that tested the prototype, but the findings were satisfactory and indicate that virtual environments can provide benefits to this kind of representation.

Geographic

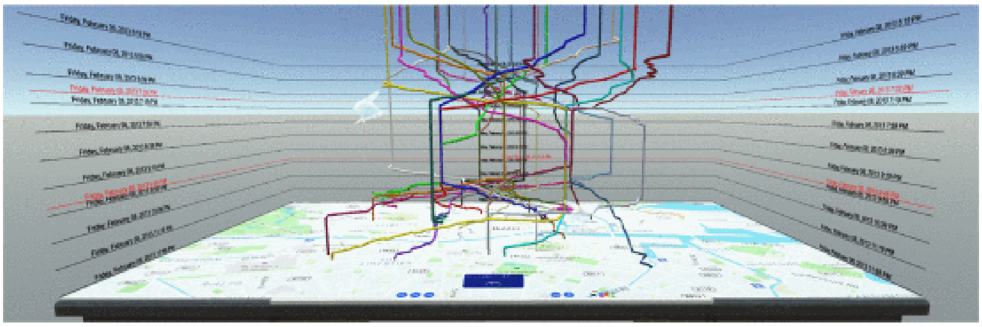


Figure 2.4: Space-Time Cube visualization of movement trajectories data in an immersive virtual environment [27]

Geographic visualization is considered the best technique for geospacial information [37], thus using map-based representations that are usually complemented with other visualization techniques. This approach is fairly explored in VR environments due to these providing a good amount of available space and navigation techniques in comparison to desktop-based setups.

Filho *et al.* [27] developed space-time cubes, or STCs, for studying spatio-temporal aspects in datasets regarding movement trajectories, while employing interaction approaches that account for visual clutter, a lack of depth information, and the general navigation issues that these visual techniques are known for. Some design choices were considered, including time growing downhill so the viewer could see the timeline intersecting the map as the current time, and color encoding for distinct timelines, which can aid pattern recognition and location identification (see Figure 2.4). The results demonstrated that the investigation of trajectories in these immersive STCs improved.

Yang *et al.* [52] proposed flow data representations in three different scenarios: on a flat display, 3D curved globes, and a novel technique called MapsLink, which uses several maps linked (see Figure 2.5). The way the height was calculated varies from being constant, proportional to distance, and proportional to flow. The comparison of different levels of flow in different map locations was tasked and the results showed in 2D flow maps, users were faster and in 3D more

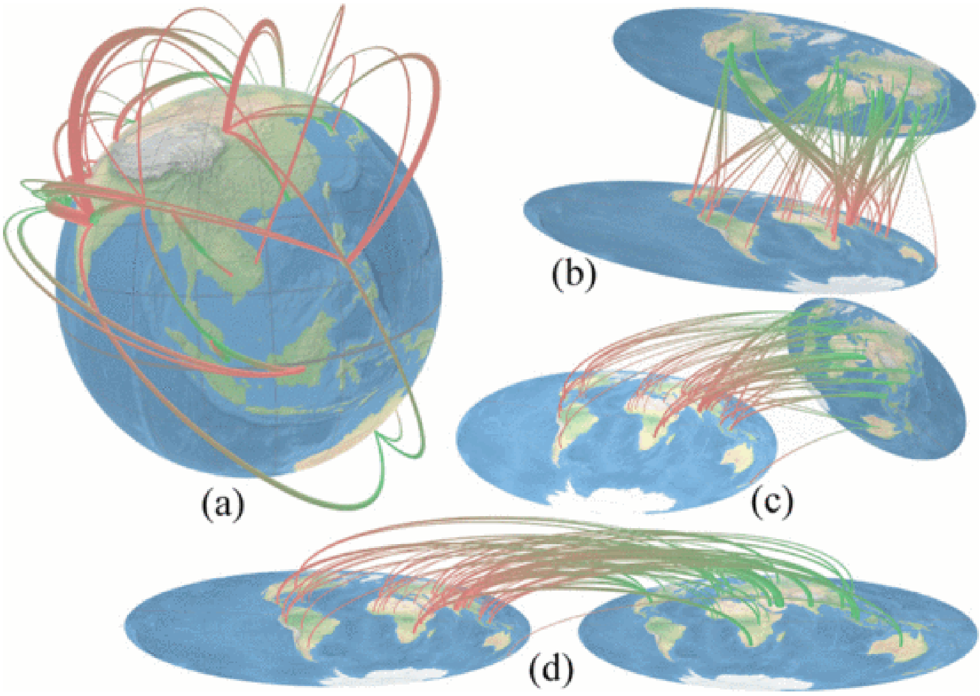


Figure 2.5: Different Movement trajectories data explored in a Flow data immersive visualization. [52]

accurate. Regarding the MapsLink, users took significantly more time performing the tasks. It was concluded that the 3D curve flow maps were the preferable approach as well as the more accurate and faster in all 3 experiences.

Glyphs, Icons and Symbols

Glyphs, icons, and symbols, as well as their benefits, are only lightly investigated in virtual worlds compared to other techniques [37]. They can, however, aid other visualization techniques in highlighting information and task execution due to their simplicity.

Gistualizer is a tool developed by Billow and Cottam [19] to visualize singular and complex data points as glyphs composed of landscapes with different properties (see Figure 2.6). Each landscape is designed when a data point is selected, and the appearance of the landscape varies depending on the data point attributes. Elements such as houses, mountains and the weather present in the visualization are coded to indicate different characteristics of the data point.

2.1.2 Visualization Tasks

The success of a visualization is mainly due to the selected technique, but also to the interaction tasks used to manage data of visibility, identifying elements, filter and general analysis of the data. In this section, we present literature related to the tasks that seem more relevant according to our objectives.

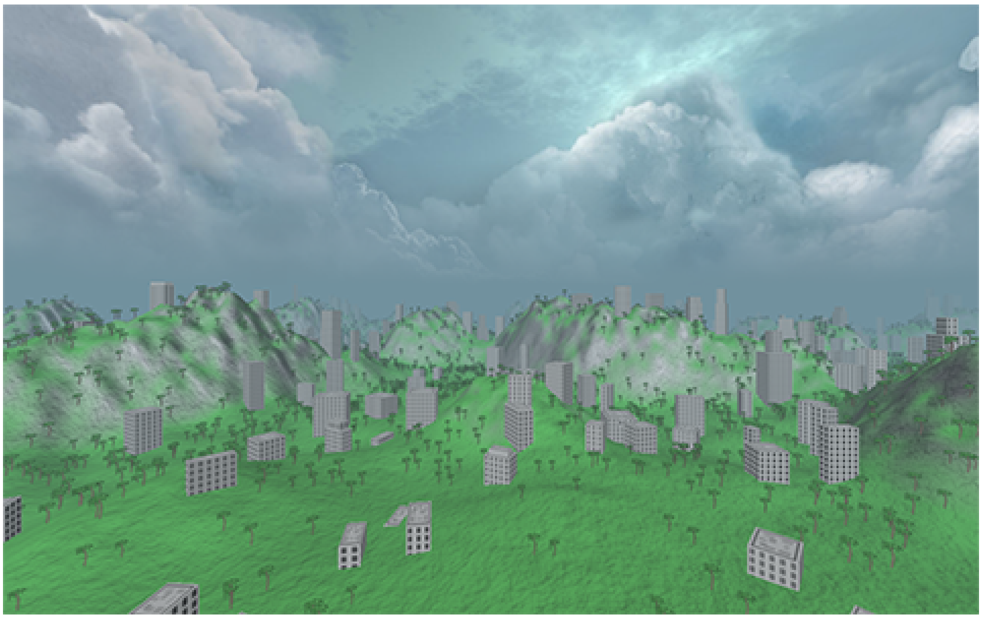


Figure 2.6: Multidimensional data representation using glyphs to display data points. Properties of the landscape correspond to different data points [19]

Clustering/Classification

Clustering and classification tasks require the user to find groups of data points that are similar [37]. This activity is commonly used to compare the effectiveness of different mediums in terms of perception and analysis. In a comparison of desktop 2D, 3D and 3D HMD VR scatterplots visualization, Filho *et al.* [26] observed that the classification tasks were statistically more accurate in the immersive environments, especially in the 3D representation, where the desktop setup did not prove to be very effective. On a similar note, Kraus *et al.* [38] compared the performance of tasks performed in 2D, 3D desktop devices, and in 3D scatterplots visualizations in VREs, being these one miniature and one full immersive representations. The findings revealed that 3D visualizations outperformed 2D representations, and that immersion can improve 3D visualizations and cluster identification.

Pattern Analysis

Pattern analysis tasks are useful, especially when analysing large data sets. Finding correlations within the data can substantially help the user understand and, possibly, reach conclusions. Raja *et al.* [44] ask the users to find trends using the same data set in 1 wall, 4-wall, with and without head tracking CAVEs. The 4 wall setup with head tracking seemed to be the preferable approach when finding the patterns. Nagel *et al.* [43] designed a system that identifies non-linear patterns in static and dynamic data sets. It also uses non-visual qualities like sound to display information in order for the user to analyse larger amounts of data at a time, which can help the detection of trends.

Visual Search

In Visual Search tasks, users are asked to find specific visual elements in a visualization that can have several different ones. These tasks can be very helpful in visualizations difficult to follow like geographic ones and, in general, their execution serves the purpose of helping the user explore the data and, possibly, reach conclusions [29], [27].

Details on Demand

Details on Demand tasks are focused on allowing a user to request more information about a specific data point without affecting the visualization as a whole, such as the current view of the representation. From a previous reference [19] developed Gistualizer and explored glyphs of landscapes, basing their properties like the mountains, the height of buildings on specific information related to a selected data point. As a result, the user can look at different aspects of the data point without interfering with the original visualization.

The visualization tasks researched above compile different ways of evaluating visualization systems, specifically in the area of data analysis.

2.2 Immersion

When it comes to analyzing large amounts of data, immersion is crucial [44]. The user can ultimately get free of the unseen limitations that more ordinary 2D interfaces inevitably provide by attempting to exploit more of the human senses. As such, Virtual Reality is the ideal medium for data analysis because it allows for the most natural engagement conceivable in comparison to other mediums. It achieves this by giving a collection of approaches for navigation, selection, text input, some of which are more immersive than others. This section is broken into sub-sections that describe what has been developed about these thus far.

2.2.1 Navigation

Teleport and continuous movement are the two types of movement that can be used for navigation (see Figure 2.7).

Teleport

Teleporting is the ability to move from one location to another instantly. The most frequent form is "Point and Teleport," in which a user with a one-hand joystick points to a destination and clicks, causing his/her location to change to the specified coordinates. Berger and Ralf [18] decided to use this navigation strategy with a clear and visual indicator of where to teleport for its low-performance requirements and already built-in implementation in the Unity3D application while testing the examination of noise standards in immersive environments. One issue they had was controlling altitude variations, which they did not limit but rather supplemented with an overview

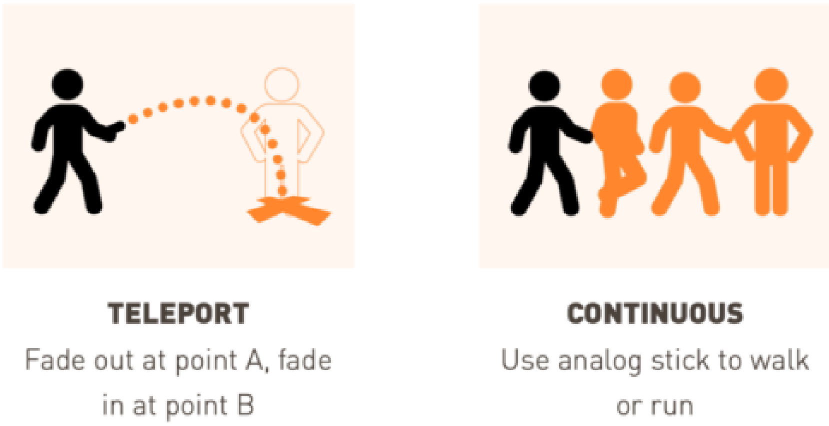


Figure 2.7: Possible navigation types explored in VRE [47]

map to ensure users would not become disoriented. Another study compared Point and Teleport to other well-known navigation approaches and found that while it is the best at avoiding collisions in a set of predefined obstacles, it falls short in terms of presenting a user with an immersive and natural experience [20].

Continuous Movement

Continuous Movement is type of locomotion in which the user progresses through the translation in a fluid motion. When compared to Teleport, the user moves in a more natural way, which results in a reduced chance of disorientation and more immersive navigation. The "Walk-in-Place" approach is a popular one that involves interpreting leg motions, while the user remains still. The way user moment was read was with the help of treadmills [28] [32], or with the use of reduced friction surfaces [30]. Walking metaphors are the opposite of the previous strategy outlined. It is based on body movements, with the exception of walking. For example, the direction the user intends to move can be detected and even the speed calculated by leaning with the user entire body or just parts of it [39], [35].

2.2.2 Selection

Objects, menu items, buttons, and different views will be selectable during the visualization. The user's interaction with the display will have a significant impact on his/her total immersion. With the advancement of virtual reality hardware, more selection approaches are being investigated in light of the hardware capabilities.

The metaphor of the user imagining his/her own hands in the virtual world is used in the virtual hand selection approach (see Figure 2.8). Actions like clicking, grabbing, stretching, rotating, and scaling visual objects will become easier to accomplish because the technique is more natural and familiar. Another option is to employ raycast pointers, which track a specific region of the user's

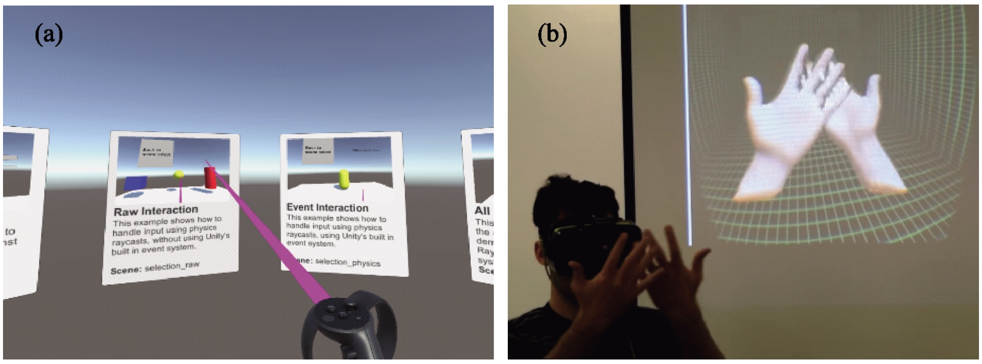


Figure 2.8: Selection Techniques in VR. a) Raycasting pointer. b) Virtual hand metaphor[31]

body and fire a focused laser at the object to be picked. Even while this method is more accurate and allows for a long-distance selection, it necessitates more complexity to perform activities other than selecting, and it might present itself as artificial to the user since it does not mimic the real world.

Wagner *et al.* [51] investigated the advantages of separating and combining these techniques in a spatio-temporal data visualization (see Figure 2.9). Despite the fact that the performance of the interaction tasks did not differ significantly, users preferred the combined method, which included a virtual hand with raycasting capabilities.

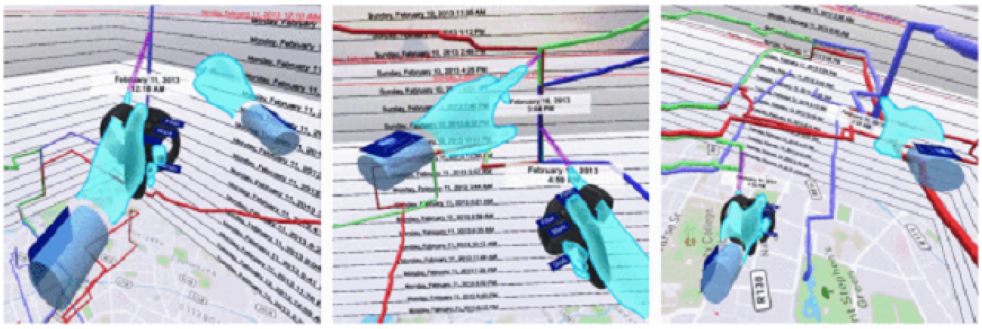


Figure 2.9: Ray pointer and virtual hand techniques combined in a geographic information visualization system [51]

2.2.3 Text Input

Although our solution does not focus on this feature, when working with computers, one of the most typical tasks is text input. In Virtual Reality, this task is not so trivial. One such approach is to create a virtual keyboard where the user can type by properly tracking which key the user wishes to press. Speicher *et al.* [46] compared six different ways of entering texting using various selection processes: a) pointing at a virtual keyboard with a projected ray from the user's head, b) using one joystick or both to be the origin of the projection, c) tapping with the controllers on the keyboard, d) and free-hand technique, which involves accurately tracking the user's hands, particularly each finger and pad-based cursor control, both discrete and continuous. The controller pointing technique demonstrated to be superior than all others but require tracked hand-held controllers.

Another option is to track the user's hand and a keyboard that is physically in front of him/her while simulating the same scenario in virtual reality. This method allows the user to write in a familiar and natural manner while simultaneously allowing him/her to visualize the moment in the virtual environment. Knierim *et al.* [36] investigate this strategy by asking 32 users to write in this precise arrangement, some of them are experienced typists and others are not. Even though inexperienced typists needed to see their avatar hands in order to write better, this study indicated that there was no substantial difficulty in the writing process.

These strategies do not deviate too much from how we normally enter text in non-immersive environments: they are just focused on writing.

Voice recognition, a natural means of expressing oneself that is very familiar to the user, may have a strong probability of challenging the text input alternatives. It is unaffected by the user's inability to visualize the real world, or by hardware limitations such as controller/hands tracking. Nevertheless, in order to work correctly, the user should be in a silent environment [33] or receive audio feedback from short range microphones. Morotti *et al.* [42] researched the benefits of visualising a fashion retail e-commerce experience, in which an expert user group was invited to try on and purchase clothing using voice commands and 3D object interaction. Although more research into this area is needed, preliminary findings reveal that the usage of voice enhanced the user's engagement and immersion in the environment.

2.3 Visualization for Finance

Information visualizations for finance support the analysis of financial relevant metrics or specific assets and, in most cases, their evolution over time. Depending on the data sets and what the user wants to depict, there are visualization techniques that are better in highlighting different relations like correlation between economic indicators, ranking of financial assets, temporal evolution and so on. As a first step to identify the most appropriate visualization technique, The Financial Times [3] identified groups of techniques that best suit different financial relationships. This way, we can observe that scatterplots are both chosen for temporal as well as correlation techniques.

As mentioned previously, current financial platforms, even though they present a news section, there is no clear explanation how these impact financial assets. Finviz [7], a website which displays financial information, allows its users to see numerous sources of information compiled in different representations like tables, maps, deviation charts with several filtering options, and also accompany worldly news updated to the minute. However, it does not cross these two topics in the same visualization environment. Bloomberg [2] is a website, which delivers several news relevant to the economic world and offers the user with tables and charts regarding stocks, futures, currency including crypto. Similar to Finviz, it does not contain visualizations that combine or visually present the correlation between the news and the financial assets charts.

Geopolitical, social, and economic events are all structured by the time dimension and, therefore, are compatible with visual representations of data across time. Timelines are an example of such representations, with the ability to depict single events that occurred over a small period of time as well as long-duration events that transcend along the time scale, which is initially set to better suit the data collection and then possibly revised by the user. SIMILE Widgets [10] provides timeline representations for a variety of events, including religious timelines, art periods, and a person's biography (see Figure 2.10). The timelines are laid up horizontally, with single events shown as symbols with a brief description and events lasting more than one time unit displayed as bars. Depending on the information they are delivering, these pieces can have different colors. This way, the user can browse over the timeline and get a basic idea of what events occurred, their types, and their duration over a specific time period. Events can have information blocks or even news sections linked to them in addition to a brief description. Knight lab [12] allows the creation of timelines with the option of further customizing how the events are shown.

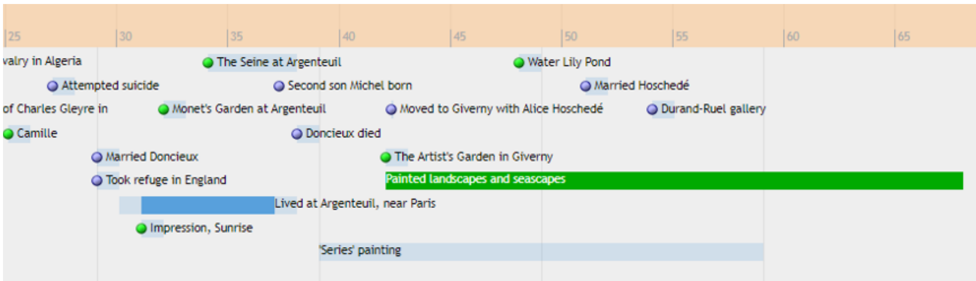


Figure 2.10: Timeline visualization of Monet's life. His life's instances are represented as circles and drawn out events as bars, both in different colours [10]

2.4 Discussion

In the preceding sections, we covered existing literature on some visualization approaches in virtual reality environments and in the finance context, as well as virtual interaction strategies used to achieve user immersion when acting in the environment. In general, we can argue that some of the strategies presented are superior to others, particularly in the domains of data visualization and finance.

Given the large complexity of financial data domains and their reliance on the variable of time, the best fit for these types of representation appears to be Scatterplots and Parallel Coordinate Plots (PCP). Scatterplots make full use of the environment's spatial capabilities, utilizing the 3D scope to accurately position data represented as points. However, in order to go beyond the three-dimensionality of the axes, certain characteristics of the plots, such as colour and size, can be used to represent additional data, which may cause visual clutter when the user is in the process of visualization.

PCPs, on the other hand, do not have this issue because it is simple to add dimensionality by reasonably increasing the number of axes employed without causing visual clutter. In addition, unlike Scatterplots, where following a single data point across several dimensions is only possible in adjacent planes, this technique allows the user to easily track it across many dimensions. When spotting trends, however, a visualization technique with the three axes connected, such as Scatterplots, may give the viewer a greater grasp of the patterns linking the data than a depiction with several but paired axes like PCPs. Glyphs, Icons and Symbols can be helpful, not as visualization technique on its own, but to complement a primary or secondary data representation.

Since our problem is related to visualizing information in a way that highlights the correlation within the data, the visualization task of recognizing patterns appears to be important to the success of the chosen visualization technique. Also, the task Details on Demand, which entails requesting additional information about a particular data point, makes sense when displaying distinct bits of data with a complex relationship. Other tasks, such as Clustering and Visual Search, have considerable advantages in our environment, but we must keep in mind that each testing session is restricted in time, which limits the activities that may be studied.

A second and auxiliary visualization can be added to the environment to offer the viewer an overall picture of the global events being considered. The timeline visualization technique can simply present a number of events and related financial indicators for the user to study and select, with the primary visualization updating dynamically.

As previously said, immersion will play a significant role in the entire visualization, assisting the user in detecting and understanding the relationships and patterns that connect the many pieces of information exhibited. There are superior interaction approaches, which help achieve naturalness and immersion in the environment. Different movement approaches might be considered depending on the size and dynamics of the display. On the one hand, the point-and-click technique, which does not require much immersion, can be useful because it is already built into most Virtual Reality softwares, especially if the display style is static and the user is largely moving in the horizontal plane. Techniques that examine walking or walking metaphors, on the other hand, expand the idea of similarity to the real world by immersing the user.

The same is true regarding selection techniques: the one that most closely resembles the real world succeeds in better immersing the user in the virtual world. Even though it is not as precise as raycast pointers, the virtual hand technique appears to be the better alternative. However, the pointing technique additionally goes above and beyond in terms of targeting and picking objects that are out of reach of the user's physical hand.

2.5 Summary

We conducted research over existing literature on the topics of Information Visualization both in Virtual Reality environments and in Finance and interaction techniques that lead to Immersion. We conclude that there are no established guidelines on which techniques, both visualization and interaction ones, to use on contextualized scenarios such as our own. This may be due to the relatively recent investment in Information Visualization techniques in Immersive Environments. Regardless, we were able to determine which visualization approaches better suit our data set domain as well as interaction methods, which lead to an improved detection and comprehension of existing patterns in the data.

Chapter 3

Immersive 2D Scatterplots and Parallel Coordinates Plots

After researching existing literature related to our work, in this chapter, we address the problem, and describe the proposed solution, detailing important aspects that are crucial to its success.

3.1 Problem Statement

The analysis of multiple sources of information is crucial when trying to reach accurate conclusions and take significant decisions. However, finding how the information is connected takes a considered amount of time, specifically when the representation does not highlight the causal relations existing in the data.

Considering the financial domain, geopolitical, social and economic events have an impact on the value of financial assets. Private investors can conduct their own research and formulate their opinion of a very wide set of sources of information. However, due to the number of updates, the volume and the heterogeneity, to read and understand all this information is a complex and time-consuming process.

Thus, a visualization including heterogeneous pieces of information designed to highlight the causal relationships within the data, could speed up the process of understanding and drawing conclusions.

3.2 Approach

We proposed the implementation of more than one visualization technique in immersive environments with several interaction methods thought to provide familiarity and immersion. As the main display, 2D Scatterplots are used to represent high-dimensional temporal data regarding the performance of assets and economic relevant markers. The users are able to interact with these plots by getting closer and selecting by touching the points that they intend to analyse.

A second visualization will dictate which plots are visible to the user. The solution includes a Timeline representing geopolitical, economic and social events in a temporal scale. The user can select which events are active, dynamically updating the Scatterplots display. For each event, several relevant financial indicators are displayed and if chosen, the respective plot will be added to the pool of Scatterplots present in the scene.

A third visualization technique, Parallel Coordinates, is also available. After selecting the intended plots, the user can opt to link the plots, better following individual points through several charts. Any selected point will also highlight points of different Scatterplots in the same time scale, including plots represented in the Parallel Coordinates visualization.

As mentioned before, interaction techniques are crucial to immerse the user into the environment and, consequently, the visualization of data. Given that these representations are not far and unreachable from the user original position, it makes sense to choose the virtual hand technique as well as physically walking as the selection and navigation techniques respectively.

Since our goal is to enable the user to find possible correlation between different sources of information, we believe a visualization encompassing multiple representation techniques and providing a natural and immersive interaction with heterogeneous data might help us more easily accomplish this goal.

3.3 Visualization Techniques

In this section, we present key aspects concerning the multiple visualization techniques integrated in our solution.

3.3.1 Timeline

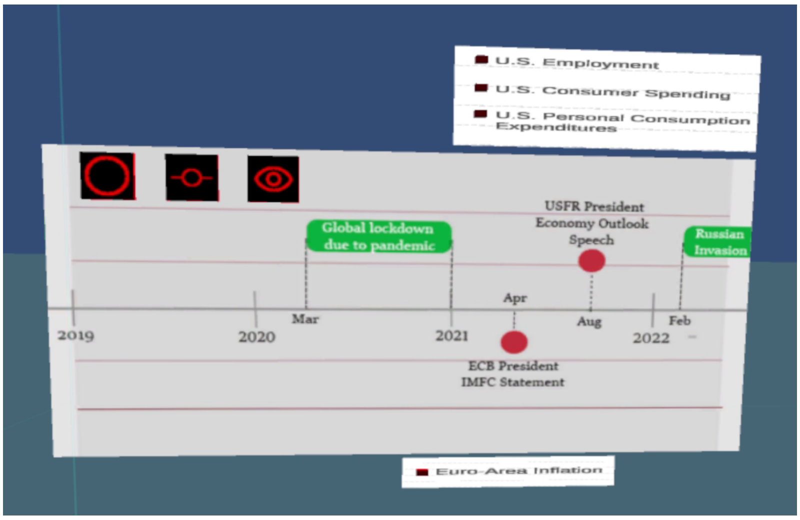


Figure 3.1: Timeline of discrete events

The Timeline visualization (Figure 3.1) establishes the relation between the several data plots and the economic, political, and social events. Along its time scale, singular and long-duration events are represented. Each event, when clicked, opens a window containing financial metrics or asset performances relevant to the chosen event. With these windows, the user can select which plots are added to the scene or close previously opened windows.

The Timeline is also responsible for managing what actions are available to the user. Aside from the event and plot selections, it has three buttons (top-left in Figure 3.1), which provide data analysis support, a different visualization technique, and to clear selected points respectively. When turning the Enhanced mode, the plot the user is looking toward gets closer to him, facilitating data analysis and further interaction. In the Parallel Coordinates mode, the active plots are cloned and positioned parallel to each other, having time corresponding points linked throughout the different plots. Lastly, there is also a button that clears any selection of points previously made.

3.3.2 Scatterplots

The 2D Scatterplots are the most common visualization method for analyzing data in our solution (example in Figure 3.2). It is composed of two orthogonal axes, where the X-axis always represents a determined period, while the Y-axis varies between monetary values, percentages, or quantity magnitudes, depending on the plot in question. Each data point is represented by a black or red sphere, depending on the time scale.

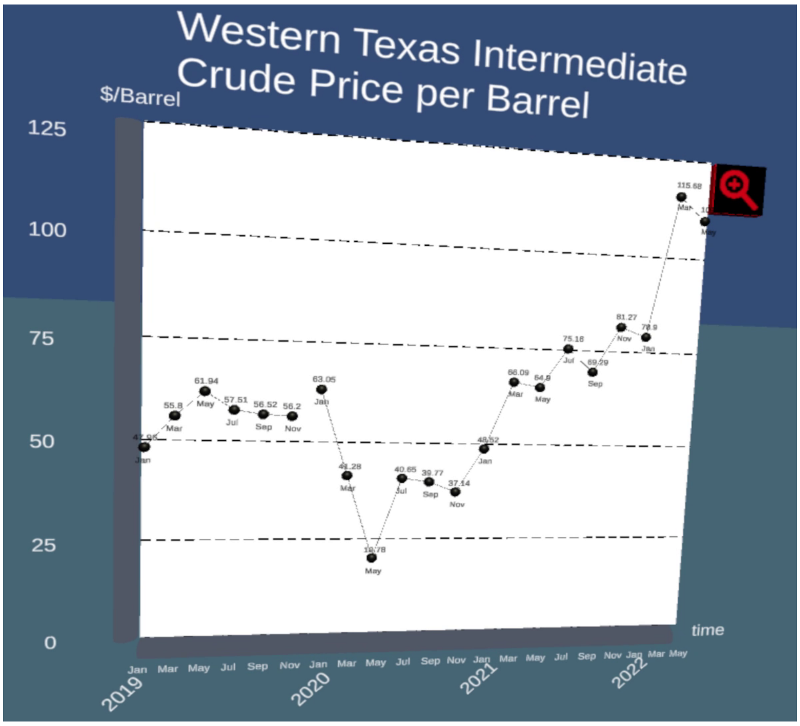


Figure 3.2: Continuous data represented by Scatterplots

For better readability, the vertical axis is divided into numerical checkpoints, easing the visualization, especially from a distance. The title, the axes measures, checkpoints, and the coordinates of the points are also scaled and appropriately placed so the user can easily understand what is in front of him.

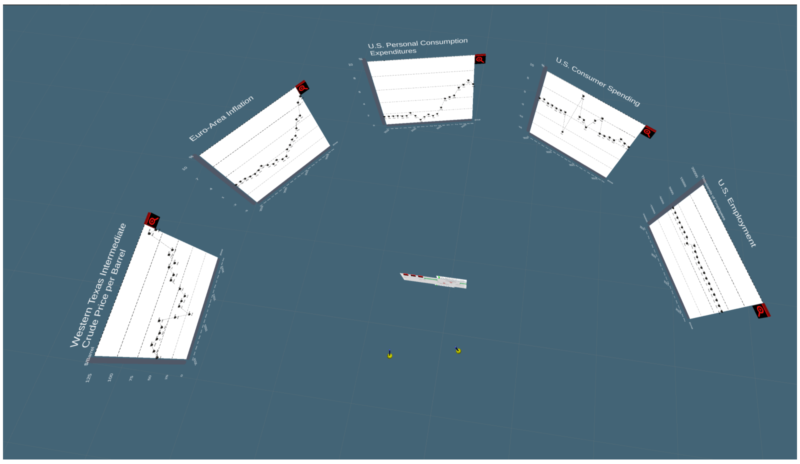


Figure 3.3: Scatterplots radial placement around the user

The plots are placed in circle around the user (see Figure 3.3), facing the center. The user can rotate the charts around 360 degrees to better position the active plots in the direction he is looking. While in Enhanced mode, the plot the user is looking towards gets closer to the center, allowing the user to get in range of selecting the points with just one or two steps. Each plot in the top right corner has the option to maximize the temporal scale, accessing more detailed information about a specific metric or asset (see Figure 3.4).

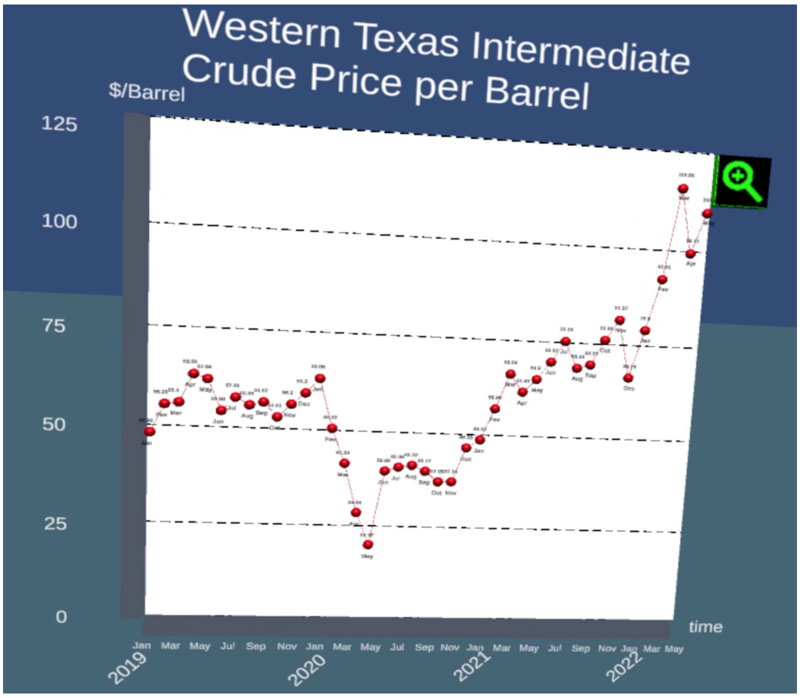


Figure 3.4: Scatterplot with its temporal scale amplified

3.3.3 Parallel Coordinates

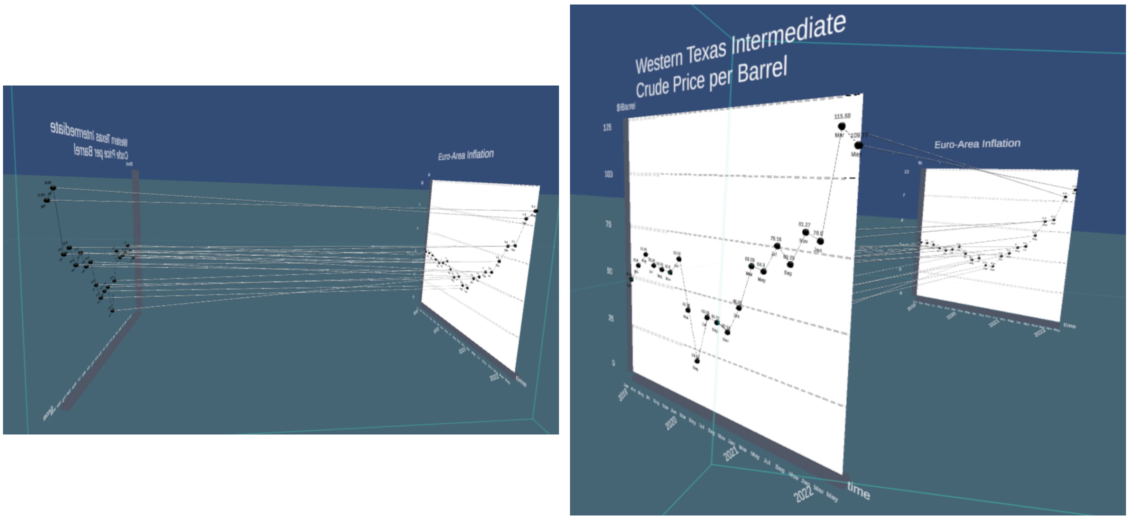


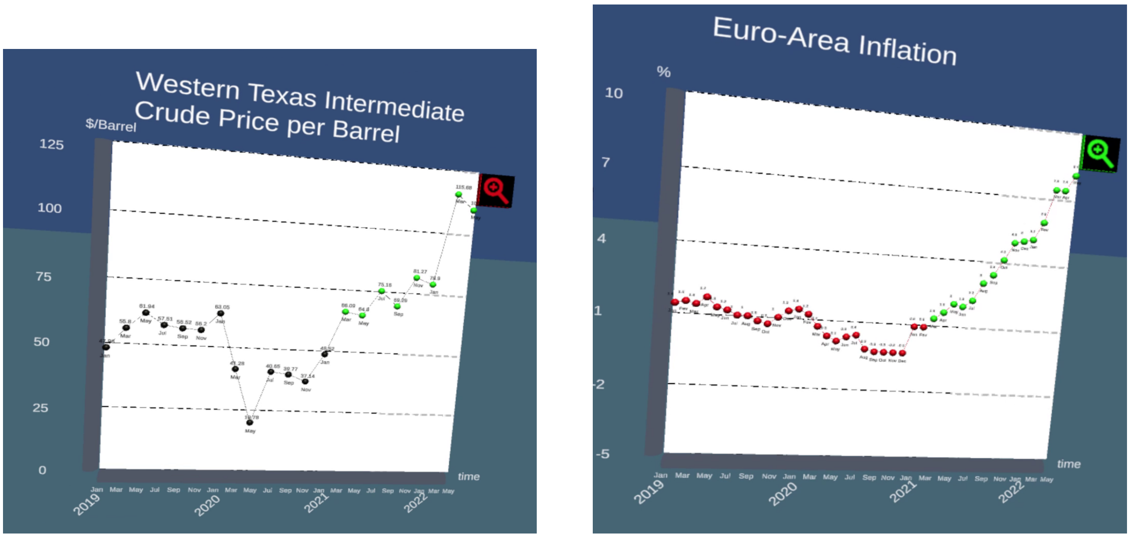
Figure 3.5: Different layouts of the Parallel Coordinates visualization

The Parallel Coordinates technique lets users quickly analyze data points across several dimensions. When the user opts for this representation, the active plots are distanced from the center and cloned. Its clones are placed in front of the user and parallel to each other. Each data point is linked to its time correspondent in the following plot, facilitating the analysis across several plots (see Figure 3.5).

The user can freely move and rotate the whole object with his/her hands. He can also scale it, but due to the almost unlimited freedom of the virtual environment, there is a minimum and maximum cap to it. If a data point is selected, consequent data points linked are also highlighted, including the link itself.

3.4 Point Selection

Point selection plays a significant role in finding possible relations between the different Scatterplot data across multiple visualization techniques. By selecting a data point from a plot, points that date the same time coordinate from different plots are also selected, highlighted with the color green (see Figure 3.6a).



(a) WTI Scatterplot with several points selected

(b) Enhanced Euro-area Inflation Scatterplot

Figure 3.6: Point Selection across different plots and time scales

When maximizing the scale, any point encompassed by the time of a reduced scale highlighted point also gets highlighted (see Figure 3.6b). These features are going to be crucial when analyzing possible correlations between plots.

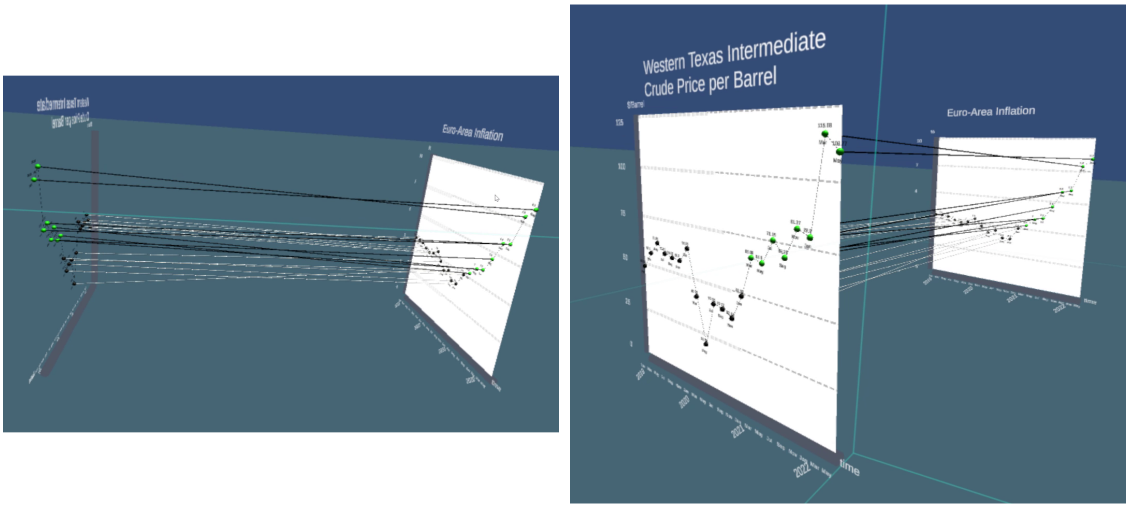


Figure 3.7: Point Selection in PCP mode

On Parallel Coordinates mode, by selecting data points, not only the point selected highlights but also the correspondent time points in neighbor plots, including the object that is linking them. This link turns black when either one of its extreme points gets selected (see Figure 3.7). Since the array of points selected is universal across both visualization methods, any selection transitions among the representations.

This interaction is performed by touching the intended spheres and reversible by repeating the action. For a better user experience, the user can clear selected points in a single operation in the Timeline.

3.5 Summary

The goal of our work is to illustrate different pieces of information in a immersive and coherent visualization that better features the existing correlations in the data. In this chapter, we present a solution, opting for visualization and immersive interaction techniques that best achieve this goal.

In our approach, we propose the implementation of Scatterplots, to represent time-dependent continuous data, Parallel Coordinates, allowing the user to follow a single data point across several dimensions, and a Timeline visualization, which acts like a menu, by giving a natural context and flow, determining which plots and visualization techniques are active. The user interaction is composed of several interaction techniques thought to be the most natural, familiar, and immersive to the user.

Chapter 4

Prototype

To adequately evaluate our proposed solution, a prototype was developed. The architecture and setup used in this prototype, as well as some implementation aspects, are described in this chapter. We also present the key features of the architecture, the essential components used in the setup, and some of the implementation choices made during the development of the prototype.

4.1 Architecture

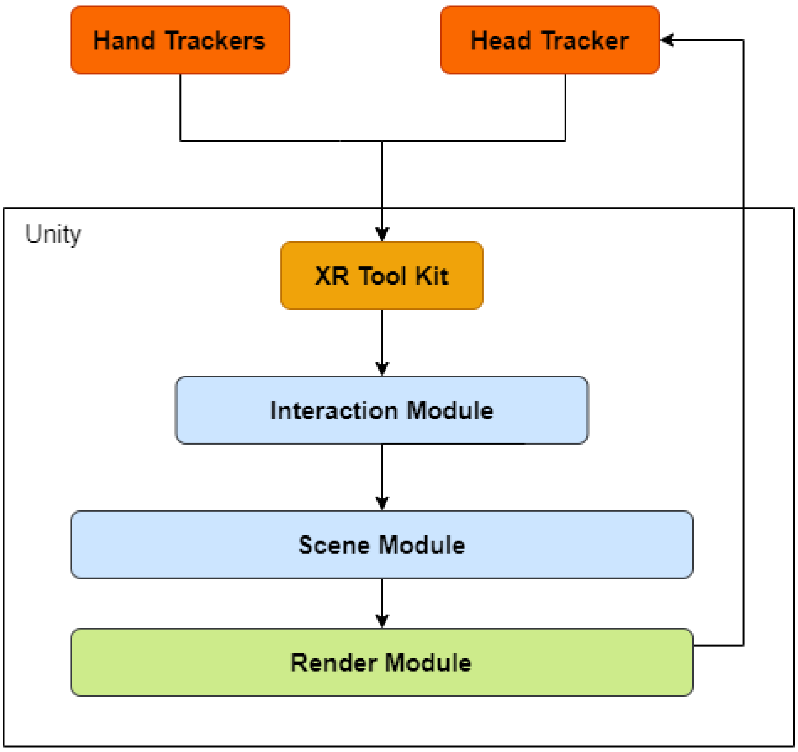


Figure 4.1: Prototype Architecture

Our prototype was built according to the architecture shown in Figure 4.1. It is composed of two trackers, each collecting data from the hardware and delivering it to our modules. The Hand

Trackers oversee the capturing all of our hand movements, including translation and rotation. It also records whether one of its input buttons is pressed, as well as the intensity with which it is pressed. The Body Tracker displays information about the user's present location. The Head Tracker informs when the user moves and/or rotates his/her head, helping to track the user's position and the direction he/she intends to look.

The XR Tool Kit package [13] is a high-level system that significantly speeds up any processing of input and movement data from the trackers used. The Interaction Module uses several XR Tool Kit functions to quickly update the user's position in the virtual environment. The Scene Module is responsible for updating the scene on the current user's position and actions. It sends the information about the scene to the Render Module, which renders the scene and displays it in the Headset.

4.2 Setup

The setup consists of several devices on the HTC Vive set, the Hand Controllers and the Headset. The prototype was delevoped in the Unity game engine, using the XR Tool Kit package to easily integrate Virtual Reality components.

4.2.1 Head Tracking

Figure 4.2: HTC VIVE Headset

For Head Tracking, we used the HTC Vive Headset (see Figure 4.2), which has 6 degrees of freedom allowing the user to look in any direction he/she wants in the 3D virtual scene, while also tracking his/her position. It has a 90Hz refresh rate and a resolution of 1080x1200 pixels per eye.

4.2.2 Hand Tracking

Figure 4.3: HTC VIVE Hand Controllers

The HTC Vive Hand Controllers (see Figure 4.3) are used for Hand Tracking, allowing users to move their hands physically while in the virtual environment. They do not provide finger tracking, although it detects 6 degrees of freedom on each user's hands. Each controller has a multifunction trackpad, grip buttons, dual-stage trigger, and system and menu buttons. We will use the trigger and grip buttons from all these input options to perform some actions in the virtual scene.

4.3 Implementation

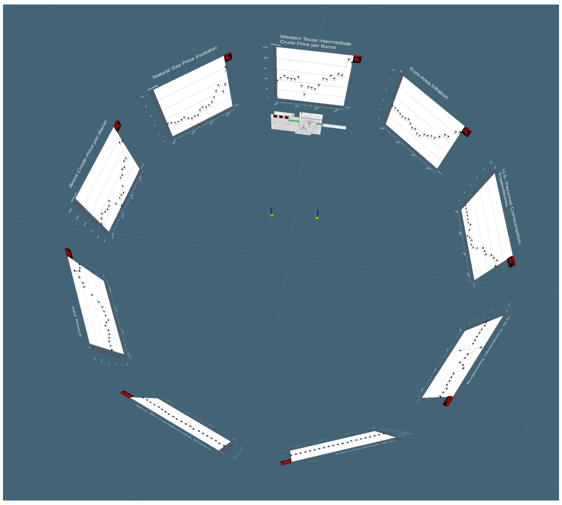
Our prototype was developed using the Unity3D engine, specifically on the 2020.3.21f1 (64bit) version with gravity and object collision disabled. We added the XR Tool Kit package that trivializes the implementation of several elementary VR interactions by automatically detecting OpenXR [14] controller inputs, such as the HTC Vive ones, and offering the hover, select and grab interactions built as components.

4.3.1 Scatterplots Design

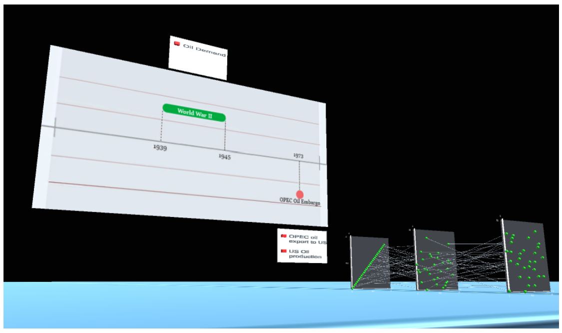
To display a dataset, we used a CSV Reader already implemented, which parses information in CSV files (see Apendix A) to a list of Dictionary data structures, having the column names as keys to access the points values. Each data point translates into a sphere, which is a prefab composed of a sphere collider and a script that changes its color according to whether its collider is triggered. The maximum size of each axis divided by the number of points per plot results in the distance between points illustrated in the Scatterplot. The axis itself is also a prefab composed of a plane as the background and two cylinders, which represent the orthogonal axes. Across these cylinders, there are several text components to display the plot measures and magnitudes checkpoints. The plane has a texture of a black stippled horizontal line over a white background. This texture is then tiled to the number of measure checkpoints requested.

When the scene starts, all the available Scatterplots are created and disabled, hidden from the user except for the primary plot the user can analyze. The other plots can be activated with the help of the Timeline. When activated, the Scatterplot in question is added to the scene adjacently to the last one. Every plot present in the scene knows of the existence of its neighbor plots, having references to the objects themselves. When a Scatterplot is removed, the object is turned off and placed in a default position, while the remaining plots are correctly positioned to fill the gap created, if any. The two functions, AddScatterplot and RemoveScatterplot (see Apendix B), are mainly responsible for placing the Scatterplot in the correct place and also updating the references on its neighbor plots.

4.3.2 Plot Placement



(a) Scaterplots radially placed around the user



(b) Parallel Coordinates distributed across the scene

Figure 4.4: Different Scatterplots' placement explored

The current solution places the Scatterplots around the user facing the center (see Figure 4.4a) while also letting the user choose a different visualization technique. However, initially, it was explored a different placement for the visualizations. In this first approach, the Timeline also dictated which plots were visible to the user, although these were always parallel to each other and connected the same way as in the Parallel Coordinates mode (see Figure 4.4b). The user had to cross large distances to analyse different plots, which could lead to loss of concentration amidst the global context.

The user physically walking is the preferable option, bringing more immersion when navigating through the representations. Therefore, another plot distribution was explored based on making the user walk smaller distances and taking advantage of the 3D space and perspective the virtual environment provides. Thus, the plots being around the user for a rapid observation of the data across the time, with the option of getting the plots closer for more detailed analysis, was a better approach than the previously explored.

Differentiating both visualization techniques, the Scatterplots and Parallel Coordinates allow a faster understanding of simple relations in the Scatterplot representations as well as the analysis of more complex analysis in the Parallel Coordinates mode.

For example, in order to see how a certain element of interest evolves, the user can quickly turn on the respective plots and analyze them from a distance. Furthermore, the Parallel Coordinates significantly speed up the response to see if a maximum occurs at the time of another maximum in different plots. The user is free to choose which visualization is best for what he wishes to understand.

4.3.3 User Interaction

The XR Tool Kit supplies several integrated interaction features and detects multiple controller properties, like input buttons and the device's current position and rotation. To change the Timeline height, the user can hold the left trigger button and move his/her hand up and down, updating the object height.

The Scatterplots displayed around the user have an empty object placed in the center as a common parent. The user, when holding the right trigger, can rotate the plots 360 degrees around him/her. With the help of the XR Tool Kit, we can measure the position of the hand controller in each frame, calculate the difference between frames and transform the result into signed Euler angles, which are added later to the vertical rotation of the plots' parent object. This way, the user can rotate the scene, ending pointing to the same point he started.

When turning on the Parallel Coordinates mode, the active plots are cloned, linked to each other, and placed in front of the user. The user can move, rotate and scale the whole object freely. Independent of which hand is used, by holding the grip button, the object parenting the Parallel Coordinates representation, turns into the respective hand object descendent. Therefore, any movement and rotation the hand object suffers is also applied to the visualization. This way, the interaction is very familiar and easy to the user, moving and rotating his/her own hand. Additionally, by holding both grips at the same time, the user can simultaneously scale all the plots. In each frame, we read the distance between both hand controllers. When both grips are held, we compare the distance to the previous frame and add the difference to the scaling property of the object parenting the whole representation.

To move the Scatterplots closer to the user in Enhanced Mode, the system takes advantage of the hover interaction [5] components the XR Tool Kit provides. The system shoots a ray from the Headset towards the direction the user is looking. This ray has a distance large enough to always collide with the Scatterplots collider, and when entering the collider, it triggers the integrated method OnHoverEntry [5], performing a function of choice afterward, which brings the plot closer to the user, allowing for better visibility of the information. The opposite occurs when the ray exits the plot collider, triggering the method OnHoverExit [5], which moves the plot to its original position. The XR Tool Kit package provides both OnHoverEntry and OnHoverExit as well as their necessary components to work.

4.3.4 Menu

The Timeline (see Figure 3.1) allows the user to select the visualization technique to analyze the information and also locks some of the interaction options available. While on Enhanced or Parallel Coordinates mode, the user cannot add, remove or rotate the Scatterplots displayed. This decision was made to avoid multiple components interacting with the properties of the same objects simultaneously.

Moreover, the selection of points is common to both visualization techniques, thus not suffering any restriction across both modes available. Point Selection, as well as the activation of the plots, is manually triggered when the user touches the object in the scene. These objects change color, giving visual feedback to the user.

4.4 Summary

In this chapter, we discussed the important elements of the architecture of the prototype system that was the development undertaken of this work, as well as its main components. We also detailed the setup and hardware used to explore the visualization and interaction techniques. Additionally, we went into detail about several of the implementation choices made during the development of the prototype.

Chapter 5

Evaluation and Results

A well-designed information visualization solution in a virtual environment may improve human understanding when analyzing the data. Determining the success of our solution is accomplished by conducting tests having real users using the system prototype and documenting their experience.

Since our goal is to assess if the Virtual Reality medium makes a difference in how information is represented and, as a result, how well it is understood, the tests consist of a comparison between the prototype usability capabilities and a Desktop baseline. The tests involve time and accuracy measurements of the users' performances of a series of tasks concerning data analysis.

5.1 Test Data

As a use case for our tests, we created an example based on the financial domain. The test data set is composed of a financial asset easily recognized and whose economic influences are very well known and documented. These influences are relevant metrics and other financial assets, whose relations are clear to the original financial asset. Furthermore, political, economic, and social events must encapsulate a time scale rich in variations regarding the value of financial assets and metrics, highlighting possible relations between them.

In total, the data set used during the lab experiments consisted of nine financial statistics, being three economic assets and six financial metrics. For each, there was a total of 40 entries, all represented in the amplified scale and half in the reduced scale. The Scatterplots represented both monthly values and every two months depending on the respective time scale.

The Western Texas Intermediate Crude price per barrel [4] is an excellent example of a financial asset complying with the above conditions. Its dependencies reach global markets, influencing several countries' economies. It has suffered several variations across the time, and several well-documented global indicators influence or are influenced by its value.

Multiple world events shape every country economy. Since the WTI Crude price influences several nations, the events chosen should make global financial indicators feel their presence.

Shortly after the start of Covid-19 pandemic, there was a global lockdown, stopping the manufacturing and production of several types of goods. This lockdown harshly impacted the demand and, consequently, production and transport of goods and ultimately, the value of Crude, making sellers willing to pay buyers to take Oil out of their storage facilities and transport vessels. China, which holds a good percentage of the world production, suffered several variations in its Manufacturing Purchasing Managers' Index and Producer Price Index [1], which provide a good read on global production and inflation, respectively.

The U.S. Employment, U.S. Consumer spending and U.S. Personal Consumption Expenditures [1] measure the country economic health, the buy of goods and services and a measure of supply and demand: inflation. These indicators provide a good overview of the U.S. enconomic status, but also the availability of products in the market. Since the United States are the top 1 Oil Producer and Oil Consumer [8], its economy is heavily influenced by the WTI Crude index value, among many other. During the Covid-19 lockdowns, the unemployment raised, people stopped buying many products and the inflation reached historical minimums.

Regarding Europe, the Euro-Area Inflation [1] is also very sensible towards crude prices, since several European countries are high oil consumers. Due to more protective legislation, the unemployment in Europe did not reach the levels seen in US lockdowns, but the citizen behaviour towards buying products was similar, leading to historical minimums in inflation during that time.

During 2020 and part of 2021 the US Federal Reserve and the European Centrall Bank performed every possible measuse to protect the economies from a higher level of degradation and made money easily available to the economies, in an attempt to avoid economies to shrink beyond manageable scennarios.

Therefore, we can consider that central currency policy has a huge impact on economy and finance. When the European Central Bank President, Christine Lagarde[1], and the Chair of the Federal Reserve of the United States, Jerome Powell[2], make speeches regarding the state of their respective economies [16] [11], the finance market listens, and immediately accomodates variations depending on the interpretation given to what they say. These speeches are a good example of the events intended, because they can disrupt periods of growth or end periods where markets are shrinking or unstable.

Hence, a considerable number of assets is directly or indirectly strongly influenced by the price of the WTI Crude (which depends on economy health, people buying products, inflation, etc.) But the same holds for other indexes such as the Brent Crude [6] and the Natural Gas [9]. Common to all these indexes is their lack of resilience towards uncertainty and "unforeseen" events. For instance, since Russia invaded Ukraine on last 24th February 2022, both assets prices skyrocketed, strongly negatively impacting countries that require to immport these materials because it is the primary way to generate energy to drive their economy. From a seller perspective, the impact may be positive if the country is benefiting from all the seel of crude and gas at higher prices.

[1] Christine Lagarde. https://en.wikipedia.org/w/index.php?title= Christine_Lagarde&oldid=1094128786, June 2022. Online; accessed 21 June 2022.

[2] Jerome Powell. https://en.wikipedia.org/w/index.php?title=Jerome_ Powell&oldid=1093400302, June 2022. Online; accessed 21 June 2022.

Given the description above, the test data encompasses a timespan between 2019 and May of 2022, which temporally contains the first Covid-19 global lockdown, the speeches of both the ECB President and USFR Chairman, and the Russian invasion, which continues in May 2022. This timeline also provides an initial interval where the user can analyze the data before the first described events started to happen.

5.2 Apparatus and Participants

The evaluation sessions took place in room I220 of FEUP (Faculdade de Engenharia da Universidade do Porto), in the GIG lab (Graphics, Interaction, and Gaming), which has limited access and so provides a relaxed and controlled environment. Users are not required to bring any piece of equipment since the room is already equipped with all the necessary testing hardware. Because of the risks involved with sharing headsets, the headset is disinfected between sessions.

The room in which the tests took part provided an ample play area for the Virtual Reality interaction. Inside the immersive environment, the users when stepping to close to the limits, would see these limits warning them of the play area ending.

A total of 20 participants (13 males and 7 females) between 16 and 35 years took part in the tests. When asked if they followed financial news or/and statistics, one answered daily, another weekly, nine every month and nine said they were indifferent. The majority had no previous experience in Virtual Reality (11 users), although 7 have used it in past projects and 2 use it weekly. Unfortunately, the users who took part in the experiments were not part of the most desired public target since they were not used to interacting with virtual environments.

5.3 Procedure

Before the test starts, users are asked to fill out a small form to declare the age, level of education, interest in financial news or statistics, and previous experience with Virtual Reality. Users also get an explanation of the goal of the experiment and of the time taken when performing four tasks in different setups, VR and Desktop.

In the VR setup, the user is allowed to experiment with the prototype for approximately ten minutes while explaining the available features the user will interact with. The Timeline is introduced, presenting to the user the sequence of actions necessary to toggle the plots and change between different visualization techniques and pointing out that amplifying the temporal scale in each Scatterplot is possible so the user can access more detailed information when necessary. After 5 or 10 minutes of experimenting with the prototype, the user can start the execution of the tasks.

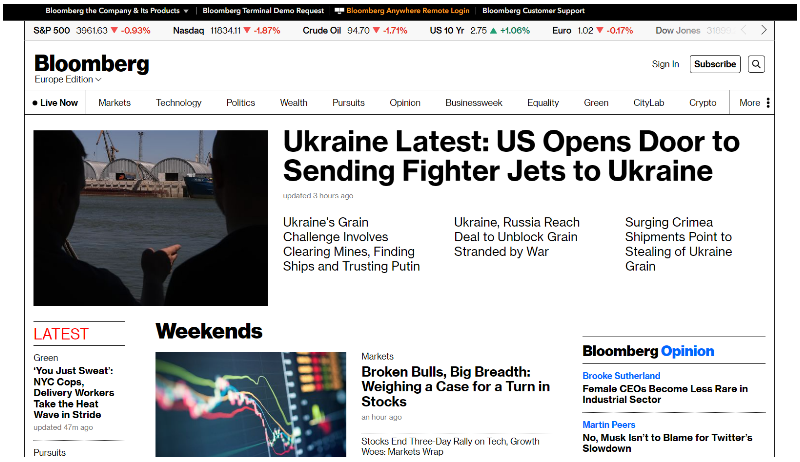


Figure 5.1: Bloomberg website front page

The Desktop setting takes place in the Bloomberg platform [2] (see Figure 5.1), which provides global financial indicators [1] and assets charts [4] [6] [9]. Every necessary tab is opened before the test starts and the users are given the opportunity to explore the multiple plots.

At the end of each setup, the user is asked to fill out another form of the System Usability Scale [21], an industry-standard composed of ten questions with five response options from Strongly Disagree to Strongly Agree. This questionnaire is a trustworthy instrument for measuring a system's usability capacities.

To avoid bias, the users would alternate between setups and corresponding tasks. Subsequent users would start in opposite setups, and the tasks shift following a Latin square design. Each testing session does not take more than half an hour to complete.

5.4 Tasks

The tests consist of the execution of four different tasks. The tasks are composed of two or three questions with financial nature promoting data analysis, whose answer is verbal. Since the financial data does not change between setups, the dates asked in each differ without changing the difficulty and the essence of the tasks.

Below, we present the tasks performed in the VR setup experimentation.

A) What was the WTI Crude Oil barrel price in **January 2020**? What was the price 6 months later, in **July 2020**?

B) Between 2019 and the present, at which date did U.S. Employment get to a **maximum**? What were the values 1 month before and after?

C) Between 2019 and the present, what is the maximum value of the WTI Crude price per barrel? In the same month, does the Euro-Area **Inflation** also reach its maximum?

D) Since the Russian Invasion that started on 24 February 2022, has the price of the barrel of the **Brent Crude Oil**, in general, increased or decreased? Did the same happen to the **Natural Gas Price**?

Additionally, we present the tasks performed in the comparison to a Desktop baseline.

A) What was the WTI Crude Oil barrel price in **March 2020**? What was the price 6 months later, in **September 2020**?

B) Between 2019 and the present, at which date did U.S. Employment get to a **minimum**? What were the values 1 month before and after?

C) Between 2019 and the present, what is the maximum value of the WTI Crude price per barrel? In the same month, does the **U.S. Personal Consumption Expenditures** also reach its maximum?

D) Since the Russian Invasion that started on 24 February 2022, has the price of the **Natural Gas**, in general, increased or decreased? Did the same happen to the price of the barrel of the **Brent Crude Oil**?

5.5 Results

During the tests, both quantitative and qualitative results were gathered. The users' time executing each task in different setups was recorded and, afterwards, they were asked to fill the standard industry questionnaire SUS, a reliable tool to evaluate systems. The success rate of each task was also assessed, considering a failed task if the user answered one of the questions wrongly.

Regarding the time and the questionnaire data, a Shapiro-Wisk Test was performed to verify the normality between setups. Then, a Paired T-test was performed to ascertain if significant differences existed in the normal distributed data. Additionally, we compare the success rate data to notice if the overall performance of each task is the same in both setups.

Concerning the user feedback, any notes the user gave during the test were collected while also providing the opportunity to write any suggestions or thoughts about the prototype and the whole experience. This data was crossed and summarized in key aspects the user referred about our solution as well as any possible improvement.

5.5.1 Time Measured

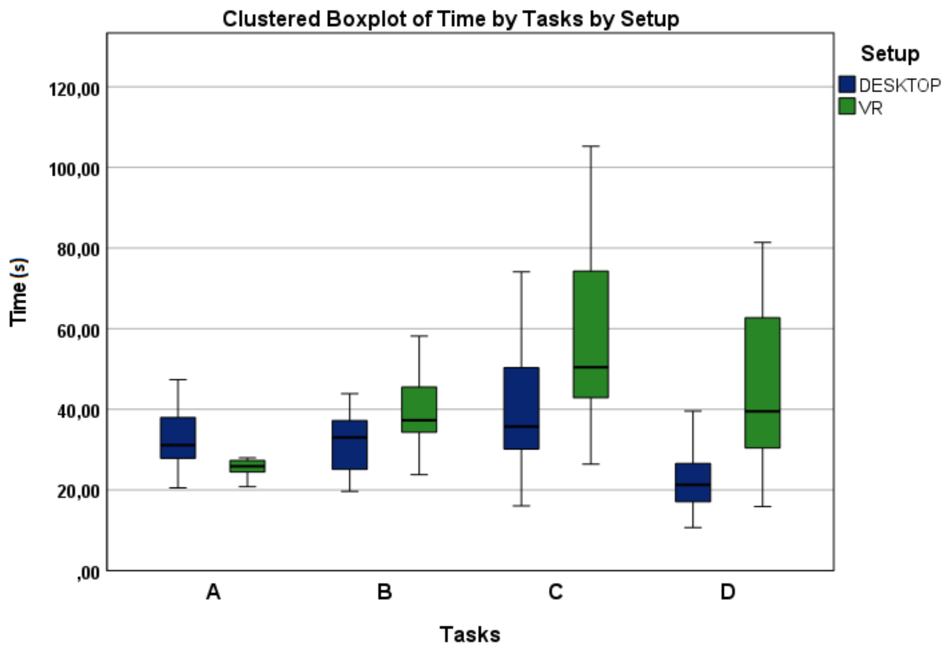


Figure 5.2: Time measured of each task performance in VR and Desktop

The total amount of time, measured in seconds, the users spent on each task is shown in Fig 5.2. The Shapiro-Wilk test was carried out and we determined the data of each task follows a normal distribution (Task A: sigVR=.223, sigDesktop: .487; Task B: sigVR=.786, sigDesktop: .544; Task C: sigVR=.077, sigDesktop: .305; Task D: sigVR=.117, sigDesktop: .345).

To verify if there was a statistically significant difference between both setups, we carried out the Paired T-test. When comparing the results of Task A ($t(12) = -3.193$, $p = 0.008$), we concluded that there was a statistically significant improvement in the time taken during the performance in VR from 31.76s ± 7.98s to 25.56s ± 1.99s ($p = 0.08$); a decline of 6.20s ± 7.00s. According to the Hedge correction, the observed effect size is high (g=.829), which indicates the magnitude between the averages of both setups.

Regarding Task B ($t(17) = 3.237$, $p = 0.005$), the results showed that the VR approach is statistically significantly slower than the Desktop, specifically from 38.162s ± 7.378s to 31.259s ± 7.052s; an increase of 6.902s ± 9.047s. Although, the effect size obtained is considered medium (g=.729).

In Task C ($t(18) = 2.982$, $p = 0.008$), similarly to Task B, the performance in the VR setup showed to be statistically significantly slower, having a total increase in time of 18.488s ± 27.028s, from 31.259s ± 7.052s to 58.198s ± 23.890s. The effect size of this test is also medium (g = .655).

Lastly, in Task D ($t(18) = 4.950$, $p < .001$), we concluded that the performance of the task on the Desktop was statistically significantly faster in comparison to the VR, from 22.468s ± 7.893s to 43.940s ± 21.268s; an increase in time of 21.472s ± 18.906s. The total effect size is high (g=0.899).

5.5.2 Questionnaire data

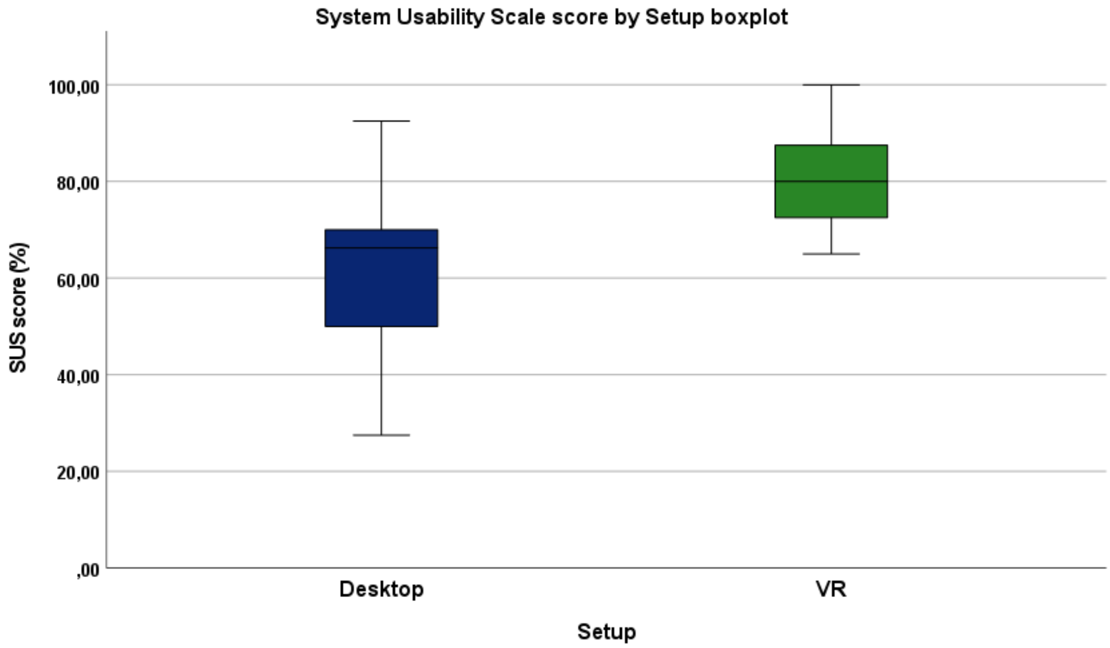


Figure 5.3: System Usability Scale scores for each of the setups tested

The System Usability Scale questionnaire can be converted to an objective result. Its score ranges from 0 to 100, with each question valued at 10 points. In odd-numbered questions, "Strongly Agree" is valued at 10 points; however, this value is zero in even-numbered questions.

The registered results are described in Fig 5.3, where we can clearly observe a difference between VR and Desktop. A Shapiro-Wilt test was carried out to determine if the users' scores followed a normal distribution, which they did (sigVR = .759, sigDesktop = .354).

Additionally, a Paired T-test was carried out to ascertain if and how much the differences between setups were in terms of statistical significance. The results obtained confirmed this relation ($t(19) = 4.331$, $p < .001$). Due to the means of the scores and the direction of the t-value, a statistically significant improvement in the VR score was concluded from 60.375 ± 17.720% to 80 ± 9.901%; an improvement of 19.625% ± 20.266%. The total effect size is high (g=.930).

5.5.3 Success Rate

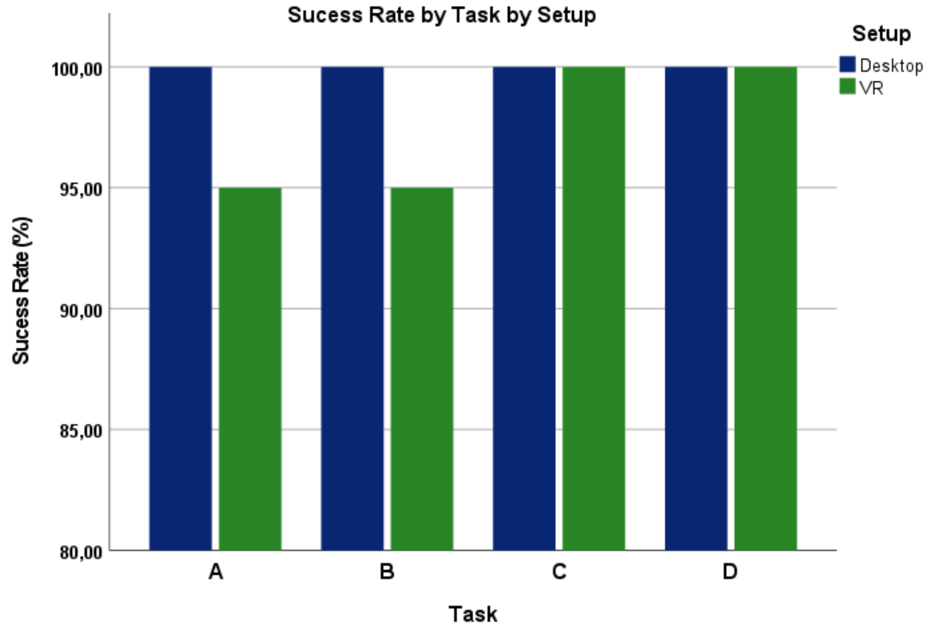


Figure 5.4: Success Rate of each task in VR and Desktop

Each task's success rate between setups was evaluated, deeming a task unsuccessful if the user provided an incorrect response to one of the questions. Fig 5.4 depicts the overall users' success rate when performing the four tasks.

Every task, which ranges from 95% to 100%, has a very high success rate in common. We concluded that all users responded correctly when testing in the Desktop setting, and only a tiny fraction of users responded incorrectly in Tasks A and B when testing in the VR environment.

5.5.4 User Feedback

During the tests, any verbal feedback as well as written suggestions or thoughts left by the users after filling out the SUS questionnaire were collected. We were able to compile several remarks about the whole experience, mainly focusing on the prototype itself.

Users felt it was intuitive and easy to learn the possible ways of interacting with the VR setup after a utility explanation of up to ten minutes. Even though describing the features was sometimes extensive, users said the interaction felt natural and immersive, and after the initial experimentation, they were confident about performing the tasks. Users who started on the Desktop compared both setups, stating that the immersive environment was a more friendly, appealing, and rapid candidate when analyzing the information. The setup's potential outside of the financial subject was also a common topic in the feedback provided.

Some improvements to the prototype were also suggested, varying from several aspects regarding the visualization techniques as well as the interaction possibilities. Users pointed out that more time could have been dedicated to the User Interface, which was rather simple. When in Enhanced mode, it was easy to lose focus on the Scatterplot when looking slightly outside its collider amidst the data analysis. In the Timeline visualization, the buttons should have labels indicating what they do. Lastly, one user also thought it was counter-intuitive to click on the Hand Controllers input buttons while immersed in the VR environment.

The feedback on the Desktop display was scarcer. Even though it didn't require much explanation before the tasks' execution, changing the tab to analyze charts located in different places on the website was a disadvantage, easily breaking the analysis context. Users also highlighted the difficulty of focusing on a particular data point due to the sensitivity of the cursor on the Desktop.

5.6 Discussion

The statistical tests to compare the setups by measuring the time of the performance of each task conclude that the Desktop approach outperformed the VR environment in three of four tasks (Task B, C, and D). The unfamiliarity with the immersive setting and the several input actions, some requiring hand controller buttons, can explain the advantage of the Desktop setup. As browsers have become the main component of our day-to-day lives, users have become used to the same interaction required during the tests. On the other hand, learning the prototype features at once before the tasks' execution can be overwhelming, making users more likely to error, delaying them. As the advantages of the VR medium get further exposed, the unfamiliarity with its properties, especially its hardware components, tends to get lower, balancing both setups in terms of prior knowledge of their interaction.

Additionally, we concluded the VR approach surpassed the Desktop setting as a usability system by comparing the data collected in the SUS questionnaire. The questions integrated into the form targeted, in a general way, both system's complexity, consistency, and intuitive and easy-to-use features. The immersive setup exceeded as a better usability system compared to the Desktop experience regarding these concepts. The design of the visualization techniques, which were heavily based on the 3D space provided by the VR medium, as well as choosing interaction features that are natural and immersive to the user, can explain the disparity between the setups.

User feedback also supported some of the findings previously stated. Overall, users pointed out that the several visualization techniques present in the prototype favored a more rapid, friendly, and appealing data analysis. They complimented the interaction with the prototype's multiple elements, saying it was intuitive, natural, and easy to learn. Some of the improvements suggested focused on the User Interface and the interaction input options. A more worked and detailed UI can further improve the features' intuitiveness and appeal. To counter the unfamiliarity with the hardware, mainly the hand controllers, the coding of hand gestures could replace the pressing of buttons outside the virtual environment. The positive feedback, as well as the improvements suggested by the users, enrich the conclusions taken and highlight the possible future work beyond this thesis.

5.7 Summary

In this chapter, we described the methodology behind the formal evaluation of the virtual solution against a Desktop baseline using real users to test. We presented several aspects of the test experiments: the use case, the apparatus and participants, the procedure during each evaluation session, the tasks the users performed, and the results collected.

After a statistical analysis of the quantitative results and the interpretation of feedback provided by the user, we concluded that our approach is more intuitive, friendly and easy to use but slower when analysing information.

Chapter 6

Conclusions and Future Work

Nowadays, people have access to a wide range of data, from high-level and general statistics to up-to-the-minute news. However, understanding how these relate to real-life events is an elaborate and time-consuming process due to its heterogeneity, volume, and number of constant updates. The role of Information Visualization in virtual environments is to immerse the user in visualizations that better highlight the causal relationships within the data and therefore provide a better understanding of the data.

We researched visualization techniques suitable for heterogeneous sources of information. We also considered interaction methods in VR environments, which are more natural and familiar, thus immersing the user in the virtual environment and the data itself.

With this in mind, we studied and developed a visual narrative encompassing more than one visualization technique. To represent time-dependent continuous data, we used 2D Scatterplots. Additionally, the Parallel Coordinates technique allows tracking a specific data point across several dimensions. The Timeline visualization connects the information to discrete events making it easier to correlate different types of information. The user interaction comprises multiple approaches thought to be the most natural, familiar, and immersive to the user.

A formal user evaluation was performed, where the VR prototype was compared to a Desktop setting. The results from the system usability questionnaire and the user feedback showed that our solution is more appealing, intuitive, and easy to use than a Desktop baseline. However, the time measured during the performance of each task determined that our solution did not provide the fastest data analysis between both setups. Furthermore, we also concluded some improvements that could enhance the prototype capabilities as future work beyond this dissertation.

Regarding the problem stated in section 1.2, we believe the developed solution may be a good alternative to visualizing heterogeneous pieces of information. Furthermore, immersing the user in the environment and, consequently, in the data helps ease the discovery of relations between data elements and reach more insight. Thus, we considered that the user would more easily understand the information and reach complex conclusions using the proposed solution.

Concerning the objectives declared in section 1.4, we met these goals by designing Scatterplots, Parallel Coordinates Plots, and Timeline visualization techniques and incorporating them satisfactorily with diversified pieces of information. We also developed an interaction model intended to be natural and immersed to the user, thus facilitating data analysis. A formal evaluation with real users composed of test experiments and questionnaires testified to the success of implementing these techniques, therefore showing the potential of virtual environments in the field of Information Visualization.

In relation to the hypothesis mentioned in section 1.3, we believe the intuitive and easy-to-use features incorporated in the prototype facilitate complex data analysis. We consider that the visualizations and interaction techniques would help the user more easily understand the information to be analyzed.

6.1 Future Work

There were certain elements in this work that have room for improvement or, alternatively, can be the ground for future work:

- **Improve the User Interface design**

  The appeal, intuitiveness, and usability of a solution can all be improved with a more polished and thorough UI. These features will make the prototype a better utility system.

- **Evaluate with more complex data analysis tasks**

  The tasks used to evaluate our prototype were rather simple. By conducting tests with more complex tasks, the solution may present better time results compared to the desktop baseline when analyzing the information.

- **Test with more experienced users**

  Users with significant virtual environment expertise can assist in overcoming the challenge of being unfamiliar with the hardware components, leading to faster data analysis.

- **Explore different input options**

  Hand gestures are worth exploring as an alternative to the hand controller input options. This way, all the interaction is visible inside the virtual environment itself, thus providing a more immersive experience.

- **Explore different data domains and visualization techniques**

  The proposed solution includes multiple visualization techniques that better suit the financial data domain. It is worth exploring different data domains and respective representation approaches to evaluate the VR medium viability for Information Visualization.

References

[1] The 12 Global Economic Indicators to Watch. https://www.bloomberg.com/graphics/world-economic-indicators-dashboard/. Online; accessed 29 May 2022.

[2] Bloomberg.com. https://www.bloomberg.com/europe. Online; accessed 23 February 2022.

[3] chart-doctor/visual-vocabulary at main · Financial-Times/chart-doctor. https://github.com/Financial-Times/chart-doctor. Online; accessed 21 February 2022.

[4] CL1 Commodity Quote - Generic 1st 'CL' Future. https://www.bloomberg.com/quote/CL1:COM. Online; accessed 21 June 2022.

[5] Class XRBaseInteractable | XR Interaction Toolkit | 0.9.4-preview. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@0.9/api/UnityEngine.XR.Interaction.Toolkit.XRBaseInteractable.html. Online; accessed 19 June 2022.

[6] CO1 Commodity Quote - Generic 1st 'CO' Future - Bloomberg Markets. https://www.bloomberg.com/quote/CO1:COM. Online; accessed 21 June 2022.

[7] FINVIZ.com - Stock Screener. https://finviz.com/. Online; accessed 21 February 2022.

[8] Frequently Asked Questions (FAQs) - U.S. Energy Information Administration (EIA). https://www.eia.gov/tools/faqs/faq.php. Online; accessed 21 June 2022.

[9] NG1 Commodity Quote - Generic 1st 'NG' Future - Bloomberg Markets. https://www.bloomberg.com/quote/NG1:COM. Online; accessed 21 June 2022.

[10] SIMILE Widgets | Timeline. http://www.simile-widgets.org/timeline/. Online; accessed 23 February 2022.

[11] Speech by Vice Chair Clarida on outlooks, outcomes, and prospects for U.S. monetary policy. https://www.federalreserve.gov/newsevents/speech/clarida20210804a.html. Online; accessed 21 June 2022.

[12] Timeline JS. https://timeline.knightlab.com/. Online; accessed 01 March 2022.

[13] XR Interaction Toolkit | XR Interaction Toolkit | 2.0.2. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.0/manual/index.html. Online; accessed 19 June 2022.

[14] OpenXR - High-performance access to AR and VR —collectively known as XR— platforms and devices. https://www.khronos.org/openxr/, December 2016. Online; accessed 19 June 2022.

[15] L. Arms, D. Cook, and C. Cruz-Neira. The benefits of statistical visualization in an immersive environment. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, pages 88–95, March 1999. ISSN: 1087-8270.

[16] European Central Bank. IMFC Statement. https://www.ecb.europa.eu/press/key/date/2021/html/ecb.sp210408~87574463be.en.html, April 2021. Online; accessed 21 June 2022.

[17] Benjamin B. Bederson, Professor Benjamin B. Bederson, Ben Shneiderman, and University Professor Ben Shneiderman. *The Craft of Information Visualization: Readings and Reflections*. Morgan Kaufmann, 2003. Google-Books-ID: IDaFnCl6TvIC.

[18] Markus Berger and Ralf Bill. Combining VR Visualization and Sonification for Immersive Exploration of Urban Noise Standards. *Multimodal Technologies and Interaction*, 3:34, May 2019.

[19] Todd V Billow and Joseph A Cottam. Exploring the Use of Heuristics for Evaluation of an Immersive Analytic System. page 5.

[20] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. Point & Teleport Locomotion Technique for Virtual Reality. CHI PLAY ’16, pages 205–216, New York, NY, USA, October 2016.

[21] John Brooke. SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, November 1995.

[22] Stuart Card. *INFORMATION VISUALIZATION*. CRC Press, March 2009. Pages: 199-234 Publication Title: Human-Computer Interaction.

[23] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST ’17, pages 71–83, New York, NY, USA, October 2017. Association for Computing Machinery.

[24] Chaouki Daassi, Laurence Nigay, and Marie-christine Fauvet. A taxonomy of temporal data visualization techniques. *Information-Interaction-Intelligence*, 2012.

[25] Ciro Donalek, S. G. Djorgovski, Scott Davidoff, Alex Cioc, Anwell Wang, Giuseppe Longo, Jeffrey S. Norris, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, and Andrew Drake. Immersive and Collaborative Data Visualization Using Virtual Reality Platforms. *2014 IEEE International Conference on Big Data (Big Data)*, pages 609–614, October 2014. arXiv: 1410.7670.

[26] Jorge A Wagner Filho, Marina F Rey, and Luciana Nedel. Immersive Analytics of Dimensionally-Reduced Data Scatterplots. *IEEE VIS 2017 Workshop on Immersive Analytics: Exploring Future Interaction and Visualization Technologies for Data Analytics*, page 5.

[27] Jorge A. Wagner Filho, Wolfgang Stuerzlinger, and Luciana Nedel. Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):514–524, January 2020. arXiv: 1908.00580.

[28] Joyce Fung, Carol L. Richards, Francine Malouin, Bradford J. McFadyen, and Anouk Lamontagne. A Treadmill and Motion Coupled Virtual Reality System for Gait Training Post-Stroke. *CyberPsychology & Behavior*, 9(2):157–162, April 2006. Publisher: Mary Ann Liebert, Inc., publishers.

[29] Carolin Helbig, Hans-Stefan Bauer, Karsten Rink, Volker Wulfmeyer, Michael Frank, and Olaf Kolditz. Concept and workflow for 3D visualization of atmospheric data in a virtual reality environment for analytical approaches. *Environmental Earth Sciences*, 72:3767–3780, November 2014.

[30] Jiungyao Huang. An omnidirectional stroll-based virtual reality interface and its application on overhead crane training. *Multimedia, IEEE Transactions on*, 5:39–51, April 2003.

[31] Susu Huang, Daqing Qi, Jiabin Yuan, and Huawei Tu. Review of studies on target acquisition in virtual reality based on the crossing paradigm. *Virtual Reality & Intelligent Hardware*, 1(3):251–264, June 2019.

[32] Hiroo Iwata. The Torus Treadmill: realizing locomotion in VEs. *IEEE Computer Graphics and Applications*, 19(6):30–35, November 1999. Conference Name: IEEE Computer Graphics and Applications.

[33] Janis G. Jimenez and Jürgen P. Schulze. Continuous-Motion Text Input in Virtual Reality. *Electronic Imaging*, 2018(3):450–1–450–6, January 2018.

[34] Daniel A Keim. Information Visualization. In *Human-Centered Issues and Perspectives*, page 48. Springer Berlin, Heidelberg.

[35] Alexandra Kitson, Abraham M. Hashemian, Ekaterina R. Stepanova, Ernst Kruijff, and Bernhard E. Riecke. Lean into it: Exploring leaning-based motion cueing interfaces for virtual reality movement. In *2017 IEEE Virtual Reality (VR)*, pages 215–216, March 2017. ISSN: 2375-5334.

[36] Pascal Knierim, Valentin Schwind, Anna Maria Feit, Florian Nieuwenhuizen, and Niels Henze. Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–9. Association for Computing Machinery, New York, NY, USA, April 2018.

[37] Matthias Kraus, Johannes Fuchs, Björn Sommer, Karsten Klein, Ulrich Engelke, Daniel Keim, and Falk Schreiber. Immersive Analytics with Abstract 3D Visualizations: A Survey. *Computer Graphics Forum*. https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14430.

[38] Matthias Kraus, Niklas Weiler, Daniela Oelke, Johannes Kehrer, Daniel A. Keim, and Johannes Fuchs. The Impact of Immersion on Cluster Identification Tasks. *IEEE transactions on visualization and computer graphics*, 26(1):525–535, January 2020.

[39] Maud Marchal, Julien Pettré, and Anatole Lécuyer. Joyman: A human-scale joystick for navigating in virtual worlds. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 19–26, March 2011.

[40] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas. *Immersive Analytics*. Springer, October 2018. Google-Books-ID: vaVyDwAAQBAJ.

[41] Bruce H. McCirmick. Visualization in scientific computing. *Computer Graphics 21*, pages 3–3, November 1987.

[42] Elena Morotti, Lorenzo Donatiello, and Gustavo Marfia. Fostering fashion retail experiences through virtual reality and voice assistants. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 338–342, March 2020.

[43] Henrik Nagel, Erik Granum, Søren Bovbjerg, and Michael Vittrup. Immersive Visual Data Mining: The 3DVDM Approach. April 2008.

[44] Dheva Raja, Doug Bowman, John Lucas, Chris North, and Virginia Tech. Exploring the benefits of immersion in abstract information visualization. *8th Int'l Immersive Projection Technology Workshop (IPT '04)*, January 2004.

[45] Ronell Sicat, Jiabao Li, JunYoung Choi, Maxime Cordeil, Won ki Jeong, Benjamin Bach, and Hanspeter Pfister. Dxr: A toolkit for building immersive data visualizations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings IEEE Information Visualization 2018)*, 25(1):715–725, 2019.

[46] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. Selection-based Text Entry in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13. Association for Computing Machinery, New York, NY, USA, April 2018.

[47] VRGear Staff. Half-Life: Alyx Won't Have Regular Teleportation. https://vrgear.com/news/half-life-alyx-wont-have-regular-teleportation/, March 2020. Online; accessed 28 February 2022.

[48] Slawomir Konrad Tadeja, Timoleon Kipouros, and Per Ola Kristensson. Exploring Parallel Coordinates Plots in Virtual Reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pages 1–6, New York, NY, USA, May 2019. Association for Computing Machinery.

[49] Bruce H. Thomas. Virtual Reality for Information Visualization Might Just Work This Time. *Frontiers in Robotics and AI*, 6, 2019.

[50] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[51] Jorge Wagner, Wolfgang Stuerzlinger, and Luciana Nedel. Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions for Data Manipulation in Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2513–2523, May 2021. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[52] Yalong Yang, Tim Dwyer, Bernhard Jenny, Kim Marriott, Maxime Cordeil, and Haohui Chen. Origin-Destination Flow Maps in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):693–703, January 2019. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

Appendix A

CSV sample file

In this section, we present a CSV sample file which was a part of the dataset used during the prototype testing phase.

a	b	c	d	e	f	g	h	i	j
2000	47,96	49,5	0,1	279	0,4	1,5	1,4	57,06	3,04
2001	55,80	50,5	0,4	224	0,9	1,5	1,4	65,07	2,86
2002	61,94	49,4	0,6	77	0,3	1,5	1,3	70,85	2,57
2003	57,51	49,7	-0,3	78	0,4	1,5	1,0	64,23	2,42
2004	56,52	49,8	-1,2	163	0,2	1,4	0,8	61,54	2,50
2005	56,20	50,2	-1,4	252	0,4	1,4	1,0	61,69	2,71
2006	63,05	50,0	0,1	339	0,6	1,9	1,4	68,60	2,13
2007	41,28	52,0	-1,5	-1498	-6,9	1,3	0,7	45,27	1,71
2008	19,78	50,6	-3,7	2642	8,6	0,5	0,1	26,44	1,89
2009	40,65	51,1	-2,4	1388	1,7	1,0	0,4	42,80	1,73
2010	39,77	51,5	-2,1	919	1,5	1,4	-0,3	42,66	2,59
2011	37,14	52,1	-1,5	333	-0,5	1,1	-0,3	39,45	2,89
2012	48,52	51,3	0,3	520	3,3	1,4	0,9	51,80	2,54
2013	66,09	51,9	4,4	704	5,2	2,5	1,3	69,36	2,70
2014	64,90	51,0	9,0	447	0,0	4,0	2,0	68,28	2,96
2015	75,16	50,4	9,0	689	0,1	4,2	2,2	76,17	3,70
2016	69,29	49,6	10,7	424	0,6	4,4	3,4	72,61	4,71
2017	81,27	50,1	12,9	647	0,5	5,6	4,9	82,74	5,52
2018	78,90	50,1	9,1	504	2,1	6,0	5,1	81,75	3,92
2019	115,68	49,5	8,3	398	1,4	6,6	7,4	118,11	5,02
2020	109,77	49,6	7,9	390	1,0	6,0	8,1	112,39	8,04

Table A.1: Dataset CSV file used in the experimentation of the prototype

Appendix B

Toggle of Scatterplot code functions

In this section, we display the code of the two functions responsible for the logic behind turning on and off Scatterplots. Their construction is similar, both start by finding the respective plot and update the information on neighbor plots depending on the user action. Afterwards, they deal with the scene changes correcting the position of the next plots if they exist.

```
public void AddScatterplot( int index) {

    //Find the last active Scatterplot
    Scatterplot last = dataPointsList[0];

    while(last.next != null) {
        last = last.next;
    }

    Scatterplot plot = dataPointsList[index];


    //Update the links between the Scatterplots
    last.next = plot;
    plot.holder.SetActive(true);
    plot.active = true;
    plot.before = last;

    //Calculating correct placement
    float angleIteration = 360 / 9;

    float currentRotation = angleIteration * activeScatterplots + PlotHolder.
        transform.localRotation.eulerAngles.y;

    plot.holder.transform.Rotate(new Vector3(0, currentRotation, 0));
    plot.holder.transform.Translate(new Vector3(0, 1, distFromPlayer));

    //Place as a son of the PlotHolder
```

```
        plot.holder.transform.parent = PlotHolder.transform;

        plot.holder.GetComponent<ChangePlot>().updateInitTransform(plot.holder.
            transform.position, camera, this);

        //Incrementing the number of active plots
        activeScatterplots++;
    }
```

```
    public void RemoveScatterplot(int index) {

        //Find the Scatterplots to be removed
        GameObject plot = GameObject.Find("Scatterplot " + index);

        plot.SetActive(false);

        Scatterplot current = dataPointsList[index];

        //Correct any gap caused by the removal of the Scatterplot
        updateScatterplotsPosition(current);

        //Update the plot to its default position
        current.holder.transform.localRotation = initRotation;
        current.holder.transform.Rotate(0, -PlotHolder.transform.localRotation.
            eulerAngles.y, 0);
        current.holder.transform.localPosition = initPosition;
        current.active = false;


        // Correct the Scatterplots links
        if(current.before != null) {
            if(current.next != null) {
                current.before.next = current.next;
                current.next.before = current.before;
            }
            else {
                current.before.next = null;
            }
        }
        else {
            if(current.next != null) {
                current.next.before = null;
            }
        }

        current.before = null;
        current.next = null;
```

```
        //Decrement the number of active plots
        activeScatterplots--;
    }
```