Virtual Reality Enabled Immersive Data Visualization for Data Analysis

Sagi Baigabulov[1] and Madina Ipalakova[1]

[1]International Information Technology University, Manas St. 34/1, Almaty, 050040, Kazakhstan

Abstract

Immersive data visualization solves the problems of traditional visualization methods, such as limited spatial sensitivity, lack of interactivity, and low level of ability to visualize complex 3D data structures, while providing immersion in virtual space, interacting with data, exploring it in a 3D environment, gaining new knowledge and patterns from data. This creates a deeper and richer data visualization experience, allowing users to discover hidden patterns and relationships between different aspects of the data that may be missed when using traditional visualization methods. There is a need for new methods and algorithms for data rendering as part of immersive data visualization in virtual environments since existing systems and solutions suffer from the following factors, such as long processing and preparation of one-dimensional and multidimensional data for rendering and analysis, inaccuracy of visualization due to excessive compression and processing, loss of necessary details, poor visual quality. This article provides a detailed overview of the process of data visualization, the difference between traditional methods and immersive methods, existing problems and solutions to these factors, and also proposes an integrated approach to immersive data visualization for the analysis of one-dimensional and multidimensional data.

Keywords

Virtual reality, immersive visualization, one-dimensional and multidimensional data, data analysis, data visualization

1. Introduction

The rise of immersive data visualization in virtual reality (VR) environments creates a significant shift in how we engage with and comprehend large, complicated data sets in the ever-evolving field of data science and visualization. This ground-breaking mix of technologies not only goes beyond conventional 2D representation but also unlocks new perspectives on comprehension, judgment, and ideation. Data will no longer be just an abstract concept; instead, it will become an experiential landscape that users can explore, manipulate, and absorb in previously unheard-of ways thanks to immersive data visualization in virtual reality. The significance of traditional data visualization techniques must also be mentioned. Since immersive visualization is built on these techniques.

For researchers, analysts, and decision-makers in fields including scientific research, medical diagnosis, industrial design, and financial analysis, the fusion of immersive data visualization and virtual reality offers a dynamic platform. This interactive and sensory journey through data allows users to walk around data sets, interact with data points as if they were actual objects, and fully immerse themselves in the knowledge they desire to grasp. It goes beyond flat charts and graphs to give this immersive experience.

2. Data visualization techniques

Data visualization is a fundamental data analysis tool that acts as a bridge between raw data and actionable insights. It transforms complex datasets into easy-to-understand graphical

DTESI 2023: Proceedings of the 8th International Conference on Digital Technologies in Education, Science and Industry, December 06–07, 2023, Almaty, Kazakhstan

baigabulov.s@gmail.com (S. Baigabulov); m.ipalakova@iitu.edu.kz (M. T. Ipalakova)

0000-0002-8700-1852 (M. T. Ipalakova)

representations such as charts, graphs, and infographics. These visualizations help analysts identify patterns, trends, outliers, and correlations in data, allowing them to make informed decisions.

Effective data visualization improves decision-making by allowing data analysts and decision-makers to quickly absorb information, identify anomalies, and communicate their findings to a wider audience. It also helps with storytelling, helping convey the narrative hidden in the data. Analyzing data without visualization can be like traveling in the dark, but with it, analysts can illuminate the path to valuable information and informed choices.

2.1. Traditional data visualization

Traditional visualization techniques are methods and approaches for visually presenting data that aim to graphically depict complex datasets, allowing users to understand patterns, relationships, and trends that may not be apparent in the raw data. Existing visualization techniques transform raw data into visual elements such as points, lines, bars, or areas to facilitate understanding and analysis. Together, these graphic elements are arranged into bar, scatter and pie charts, histograms, graphs, and tables. Examples of commonly used charts are Gantt, Pareto, Venn, etc.

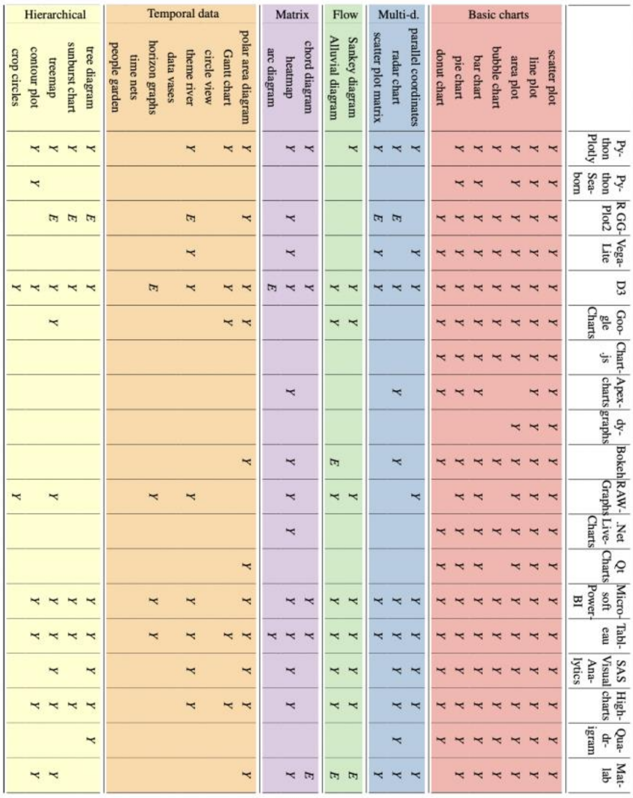
Traditional visualization techniques are used to efficiently explore, analyze, and communicate data. This type of visualization is based on human perception: a person can be visual, auditory, or kinesthetic. It is visual information that can be quite easily recreated and depicted from a large number of different data. A person, upon seeing a visual display of data, begins to match attributes of the data with visual properties such as position, length, color, or shape, using human perceptual abilities to interpret and understand the information. It is due to this factor that a person receives implicit or unusual information that helps him make informed decisions and identify patterns or anomalies.

Traditional visualization methods have been used for a long time in various fields of science and applied fields [1]:

- Business and Finance [2]: Visualization techniques are widely used in business intelligence, financial analysis and reporting. They allow you to display sales data, market trends, financial metrics, and other business elements for decision-making and strategic planning.
- Data Science and Analytics [3]: Visualization plays a critical role in data exploration, data mining, and predictive analytics. This helps analysts understand complex data structures and identify discrepancies between actual and predicted results.
- Manufacturing and Industry [4]: Visualization techniques find applications in production monitoring, quality control, and process optimization in industry. They provide real-time visual feedback on production operations, allowing operators to detect anomalies, and failures early, monitor performance and make timely adjustments.
- Healthcare and medicine [5]: Visualization of medical data, such as patient records, test results, or clinical trials, helps in diagnosis, treatment planning, and research into new or poorly understood diseases. Imaging techniques help understand medical trends, correlations, and treatment outcomes.
- Social Sciences and Humanities [6]: Visualization techniques help in social media exploration, sentiment analysis, text mining, and historical data analysis. They help researchers understand patterns, relationships, and cultural phenomena.

Traditional imaging methods have their advantages. However, over time, existing visualization methods have become worse at displaying large and complex data structures. For this reason, certain disadvantages of traditional approaches to data presentation can be identified. These include limited interactivity or its absence, inaccuracies in the context of the information provided, difficulty in conveying time attributes, etc. For example, users are limited in their ability to manipulate data and cannot dynamically interact with visualizations and explore them in real-time. When working with large amounts of data, these methods can become ineffective, resulting in increased processing time and reduced performance. Deep understanding and exploration of complex data require a sense of immersion in the data, as well as a sense of context, that traditional methods struggle to achieve when confined to two-dimensional charts and graphs. In particular, visualization of multidimensional data becomes a difficult task.

It is important to note that although traditional visualization methods have their limitations, these visualization techniques and methods provide the basis for data exploration and analysis and continue to be widely used alongside new technologies and advanced visualization approaches. However, advances in technology and the emergence of immersive visualization techniques such as VR and AR aim to address the above shortcomings and provide more interactive, immersive, and context-rich data visualizations for data analysis.



		Python Plotly	Python Seaborn	R GGPlot2	Vega-Lite	D3	Google Charts	Chart.js	Apexcharts	dygraphs	Bokeh	RAWGraphs	.Net LiveCharts	Qt Charts	Microsoft Power-BI	Tableau	SAS Visual Analytics	Highcharts	Quadrigram	Matlab
Basic charts	scatter plot	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	line plot	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	area plot	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	bubble chart	Y		Y	Y	Y	Y	Y			Y		Y		Y	Y	Y	Y	Y	Y
	bar chart	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	pie chart	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	donut chart			Y	Y	Y	Y	Y	Y		Y		Y	Y	Y	Y	Y	Y	Y	
Multi-d.	parallel coordinates	Y			Y	Y						Y			Y	Y	Y	Y		Y
	radar chart	Y		E		Y			Y		Y				Y	Y	Y	Y	Y	Y
	scatter plot matrix	Y		E	Y	Y									Y	Y				Y
Flow	Sankey diagram	Y				Y	Y					Y			Y	Y	Y	Y		E
	Alluvial diagram					Y	Y				E	Y			Y	Y	Y	Y		E
Matrix	chord diagram	Y				Y									Y	Y				E
	heatmap	Y		Y	Y	Y			Y		Y	Y	Y		Y	Y	Y	Y		Y
	arc diagram					E										Y				
Temporal data	polar area diagram	Y		Y		Y	Y				Y			Y	Y	Y	Y	Y		Y
	Gantt chart	Y				Y	Y									Y		Y		
	circle view																			
	theme river	Y		E	Y	Y						Y			Y	Y	Y	Y		
	data vases																			
	horizon graphs					E						Y			Y	Y				
	time nets																			
	people garden																			
Hierarchical	tree diagram	Y		E		Y									Y	Y	Y	Y	Y	
	sunburst chart	Y		E		Y									Y	Y		Y		
	treemap	Y		E		Y	Y					Y			Y	Y	Y	Y		Y
	contour plot	Y	Y			Y									Y	Y		Y		Y
	crop circles					Y						Y								

Figure 1: Recommended imaging techniques given in article [3]

2.2. Immersive data visualization

Immersive data visualization is a collection of data visualization techniques and approaches using technologies such as virtual reality (VR) and augmented reality (AR) to create interactivity and immersive experiences for data exploration and analysis. Immersive visualization techniques involve creating simulated environments that can reproduce real or abstract data. Users can interact with this environment using specialized hardware, such as virtual reality headsets or augmented reality devices, to visually and spatially explore and manipulate data.

Immersive data visualization solves several problems faced by traditional data visualization approaches. For example, thanks to 3D space, it is easy to visualize multidimensional data, collect the necessary details into context, and use an interactive environment for dynamic data processing.

Currently, immersive visualization methods are already widely used in various fields of activity. This approach is gradually replacing traditional visualization solutions, which were written in Chapter 2.1.

2.2.1. Virtual reality for immersive data visualization

It is worth noting that immersive visualization is based on technologies such as virtual reality (VR) and augmented reality (AR). Virtual reality (VR) is a key component of immersive visualization techniques. It involves creating a computer environment that simulates the real or imaginary world. Virtual reality headsets and controllers allow users to interact with and navigate a virtual environment. In the context of data visualization, virtual reality can provide a fully immersive experience where users can visualize and manipulate data in three dimensions, improving their spatial understanding and opening up new perspectives on the data.

These frameworks are used to create immersive visualization based on virtual reality: Unity3D, Unreal Engine, and A-Frame. Unity3D is a popular game development engine that supports virtual reality development and provides tools for creating interactive and immersive games. The Unreal Engine, like the Unity3D game engine, is widely used to create virtual reality applications, offering a number of features for immersive environments. Unlike Unity3D and Unreal Engine, A-Frame is an open-source web framework. It is used to create VR environments using HTML, CSS, and JavaScript. Because these technologies are easy to learn by less experienced developers, the development process is greatly simplified, and integration with web-based data visualization libraries is seamless.

Immersive visualization techniques with VR environments can change the way we visualize and understand data, offering new opportunities for analysis, exploration, and decision-making in various fields.

3. Implementation of immersive data visualization

In addition to the previously mentioned tools such as Unity3D, Unreal Engine, and A-Frame, other technologies make it possible to create high-quality VR environments for immersive visualization, such as:

1. Blender is an open-source 3D rendering package that can be used to create immersive data visualizations. It offers a wide range of features for modeling, animation, and rendering.
2. D3.js is a popular JavaScript library for creating dynamic and interactive visualizations of web data as web resources. Although it is primarily focused on 2D visualizations, it can be combined with VR/AR platforms to visualize data in immersive environments.
3. Vispy is a high-performance data visualization library in Python. It provides a flexible framework for creating interactive visualizations with an emphasis on GPU-accelerated rendering. Based on the OpenGL library. Introduces rendering capabilities for both 2D and 3D elements.
4. Babylon.js is a rich JavaScript framework for creating 3D games and immersive web applications. It offers support for VR and AR, making it suitable for developing immersive data visualization applications.
5. Three.js is a lightweight and versatile JavaScript library for creating web-based 3D visualizations and has a similar structure to Babylon.js. It provides an abstraction layer for WebGL.

It is worth noting that the choice of tools and frameworks depends on specific requirements, programming languages, and platforms. For example, the implementation of immersive visualization within web resources, i.e. creating a Web View is easier to do using D3.js, Babylon.js, and Three.js. Desktop applications can use Vispy since the Python language is a cross-platform solution.

3.1. Implementation of immersive data visualization in companies

High-tech companies are eagerly adopting immersive visualization tools, benefiting from a more detailed experience of analyzing data from different angles. More specifically, the list of such companies includes NVIDIA, Siemens, Microsoft, and Google.

NVIDIA, a leading technology company, has been involved in the development of immersive data visualization solutions. Their VRWorks toolkit provides a variety of tools and libraries for creating immersive virtual reality experiences, including data visualization applications.

Siemens is a multinational conglomerate. Successfully implements immersive data visualization for various applications. They use virtual reality technology to visualize complex manufacturing processes, simulate factory floor plans, and improve production efficiency.

Microsoft, as an IT giant with a wide range of implemented products, has succeeded in AR/VR. The company has developed mixed reality platforms such as HoloLens, which combine the capabilities of virtual and augmented reality. They are used in a variety of industries, including data visualization for scientific research, design, and engineering.

Google has invested in immersive data visualization through projects such as Google Earth VR, which allows users to explore geographic and spatial data in a virtual environment. They have also integrated virtual reality capabilities into their data visualization platform Google Data Studio.

Thus, it is clear that there is a positive trend toward the use of data visualization in AR/VR environments for in-depth analysis and the ability to find patterns, insights, and information hidden at first glance.

4. Challenges of immersive visualization

Immersive data visualization for data analysis is a large and resource-intensive process. Due to these factors, existing tools experience various difficulties, which in turn lead to incorrect data visualization and analysis, which are the main purposes of using these tools.

Immersive visualization requires rendering complex 3D scenes in real-time, which can be computationally intensive. Ensuring a smooth and responsive experience while maintaining high-quality graphics can be challenging. Optimizing rendering techniques, reducing latency and efficient use of hardware resources are ongoing areas of research, e.g. the task of performance and optimization always remains a priority.

Designing intuitive and effective interactions in immersive environments is critical to user engagement. Creating natural and seamless experiences that meet user expectations can be challenging. Balancing functionality, usability, and immersion while minimizing user fatigue or discomfort requires careful design, as design and user experience weigh heavily in the overall process.

As data, structures, and presentation become more complex, the process of immersive data visualization itself becomes more complex. This process involves transforming data into meaningful visual representations in a virtual environment. Selecting appropriate visualization techniques that effectively convey complex information in 3D space can be challenging. Integrating different data sources and providing real-time updates within immersive visualization remains a challenge. Processing large datasets, streaming data, and maintaining synchronization between the virtual environment and underlying data sources pose challenges. Efficient data processing, streaming, and integration techniques are essential to provide up-to-date and accurate visualizations.

In an immersive environment like VR, the user is always in the process of interacting with the environment. Because of this, calibration and tracking of user gestures and movements play an important role since correct immersive visualization often depends on these tasks. Calibrating and maintaining accurate tracking systems can be challenging, especially in dynamic environments or with complex setups. Achieving accurate and reliable tracking is critical to a seamless experience and maintaining a sense of immersion.

When implementing new visualization methods, developers of such systems are often faced with the problem of implementing different ideas, creating content, and using scalability, since creating immersive content for visualization often requires special skills and tools. Developing scalable content creation pipelines that enable the efficient production, modification, and deployment of immersive applications is also a challenge. We must strive to balance the need for richness and vibrancy of content with its scalability.

The described problems require constant development to improve the capabilities and usability of immersive visualization tools, environments, and algorithms. Collaboration among researchers, developers, and users is necessary to address these challenges and unlock the full potential of immersive visualization across domains.

4.1. Key challenges in immersive visualization

The above challenges in visualizing data for analysis certainly require attention, but challenges related to hardware, distributed and parallel processing, and rendering optimization are global challenges. Current shortcomings in the implementation of the solution are indicated in Table 1.

Table 1
Key Issues in Immersive Visualization

№	Problem domain	Key issues	Description of problem
1	Hardware	Resources - cost and availability	Lack of hardware, including virtual reality headsets, graphics processors, and systems for tracking user movements and state. The cost of purchasing and maintaining such equipment can be prohibitive, limiting availability for individuals or organizations on a limited budget.
		Compatibility and Integration	Difficulty in ensuring compatibility and seamless integration of different hardware components. Different hardware manufacturers may have their own technologies and software interfaces that require careful coordination and development efforts to ensure interoperability.
		Scalable and upgradeable	Functional limitations in hardware scalability and upgradeability to handle growing data volumes and increasing complexity. Balancing performance requirements, future scalability, and cost considerations can be challenging.
2	Data distribution and parallel processing	Data Volume and Bandwidth	Difficulty in efficiently distributing and processing data across multiple devices. Limited bandwidth or network latency may prevent real-time data from streaming and syncing across multiple users or locations.
		Load balancing and scalability	The difficulty of working with parallel data processing to balance the load of computing tasks between several processors or nodes. Achieve scalability while maintaining consistent performance across all nodes.
		Synchronization and integrity	The criticality of ensuring data integrity and synchronization between distributed processors or nodes. Delayed or inconsistent data updates can lead to visual artifacts, misconceptions, or loss of immersion.
3	Rendering optimization	Real-time rendering	The challenge of rendering 3D scenes in real time: high frame rates, smooth motion and high-quality images. Constant attention is paid to the balance of realism and performance.

Element visibility	The challenge is to minimize unnecessary computation and improve rendering performance for greater, clearer visibility.
Level of Detail (LOD)	Difficulty eliminating unnecessary details and displaying critical datasets.

Solving these problems often requires a combination of hardware improvements, algorithmic optimization, and system-level modifications. Researchers and software developers related to VR/AR technologies for immersive data visualization are constantly working to find innovative solutions to overcome these challenges and improve the overall immersive visualization experience.

5. Literature review

Table 2 provides a literature review summary of 11 related works and articles in which immersive visualization and its implementation are discussed.

Huang et al. conducted a plant ecosystem simulation using a parallel processing computer cluster equipped with conventional graphics cards. The goal of the study was to achieve more efficient modeling and visualization of large groups of plants compared to traditional methods. By implementing a scalable architecture, the researchers developed a system capable of simulating complex plant ecosystems in significantly less time. To cope with the complexity of the problem, the authors used multi-level models to simplify the simulation by dividing it into several parts that could be simulated simultaneously. The entire modeling process is subsequently visualized locally and displayed on a large immersive wall, providing immersive visualization with near real-time response. The authors' virtual plant ecosystem is specifically designed to facilitate parallel modeling and visualization of large plant ecosystems. It serves as a framework that allows for seamless integration of various simulation modules, complemented by an immersive display system. Although the current system optimization is focused on PC clusters, the authors plan to adapt it for grid platforms that provide increased computing power at a reasonable cost. Additionally, an important future goal for researchers is to incorporate more complex ecosystem simulation models into their framework [7].

Large-scale, high-resolution displays are increasingly used in next-generation interactive 3D graphics applications such as large-scale data visualization, immersive virtual environments, and participatory design. These applications require the inclusion of a highly efficient and scalable 3D rendering engine to produce high-resolution images in real time. Chen et al. in their research are currently exploring the possibility of building such a system using cost-effective standard components in a PC cluster. The focus is on developing scalable algorithms that can efficiently partition and distribute rendering tasks while taking into account the constraints imposed by the bandwidth, processing capabilities, and memory capacity of a distributed system. In this paper, the researchers compare three different approaches that differ in the nature of the data passed from the client to the display servers: control data, primitive data, or pixel data. Each approach is accompanied by a description of initial experiments conducted using a functioning prototype system to control a video wall with multiple projectors and a PC cluster. The results show that the suitability of each approach depends on the system architecture, taking into account factors such as communication bandwidth, storage capacity, and processing power available on both the client and server side of the display [8].

In our ever-expanding technological society, huge amounts of data are generated daily. The sheer volume of this data poses a challenge in understanding and using it for decision making. This problem extends not only to businesses, policymakers and scientists, but also to the general public, as data has become a valuable commodity, closely linked to our ability as data owners to extract meaningful information from it. Traditionally, this process has been complex and accessible primarily to experienced data scientists who are familiar with the intricacies of information extraction and the specialized tools used for these purposes. However, recent advances in virtual reality (VR) technology, especially with the advent of affordable hardware such as Head Mounted Displays (HMDs), have paved the way for a new area of research known as immersive analytics. This field aims to explore the potential of analytical thinking using immersive computer interfaces, studying how people perceive and interact with representations that look like real objects. This thesis, led by researcher Marius Nicolae Varga, focuses on representing multidimensional data using simple 3D geometric shapes called geons, which together form complex visual objects or glyphs presented to users in an immersive virtual environment. To construct these glyphs, a set of rules was developed based on the fundamental principles of an object recognition theory known as Component Recognition. The researcher developed a toolkit capable of representing multidimensional data sets in immersive virtual environments using a human-centered approach as a basis. Particular attention was paid to the immersive aspects of the application, with a focus on spatial immersion, data embodiment, multi-sensory presentation and immersive storytelling. A series of experiments were conducted to evaluate the effectiveness of this approach, including an evaluation of the immersive experience itself. The results show that applying structural object recognition theories to the construction of complex visual objects can facilitate the search for optimal solutions in large data sets, even for users without data mining experience. The results also highlight the significant contribution of the immersive aspect of the app to the comprehension process, with participants providing positive feedback on the level of immersion achieved [9].

Flatken et al. present the software architecture and framework that has been developed over the past decade. The main goal of this architecture and framework is to facilitate the creation of scalable and highly interactive visualizations for processing large data sets and accommodating displays of various sizes. By integrating distributed processing, data streaming, and dynamic scheduling, the platform enables view-dependent feature extraction and progressive data streaming. In addition, special attention was paid to ensure that the platform can support visualization on a variety of devices, ranging from local workstations to large virtual environments with multiple displays [10].

Cedilnik et al. use remote parallel computing resources to run scientific simulations aimed at simulating various scientific phenomena. These simulations generate large data sets, requiring the use of visualization tools for understanding. In this context, several challenges need to be addressed in order to develop an effective visualization tool capable of representing these data sets. These challenges include efficient processing and display of massive data sets, as well as seamless transfer of data and control information between geographically dispersed computing and visualization resources. The authors propose a solution based on combining a parallel data server, a parallel data rendering server and a client controller. Building on this foundation, the paper describes a wide range of integrated solutions for remote/distributed imaging challenges. These solutions include the introduction of an efficient parallel M-N algorithm for transmitting geometric data, the creation of an efficient server interface abstraction, and the implementation of parallel rendering techniques adapted to various rendering modalities, including tiled video walls and CAVEs [11].

The authors of [12] believe that one of the notable problems in the field of visualization is related to the selection of a suitable tool for conducting research projects or experiments. The field of immersive analytics (IA) shares this challenge, but has found support in game engines and web technologies to develop its own solutions, frameworks, and toolkits. While these technologies effectively address issues such as rendering and interaction, they lack the necessary functionality to facilitate data analysis in immersive environments. The authors of this article introduce ImmVis, an innovative open-source framework that allows IA applications to leverage the data analysis capabilities of well-established Python programming language libraries. The platform is designed to be compatible with a variety of platforms and programming languages, extending the capabilities of existing IA tools to offer more sophisticated data analytics capabilities.

Mobile virtual reality (VR) offers both mobility and immersion, making it an ideal platform for visualizing disaster scenes in three-dimensional (3D) format. Compared with other methods, in this context, users can perceive and recognize disaster conditions more effectively. However, achieving a high scene rendering frame rate is critical to maintaining immersion and preventing user discomfort. Current visualization approaches do not provide a satisfactory solution to this requirement. Kaloian Petkov's research [13] focuses on creating and optimizing 3D disaster scenes specifically designed for mobile virtual reality to meet stringent frame rate standards. First, a plug-in-free browser/server (B/S) architecture is designed to create and render 3D disaster scenes in mobile virtual reality. Secondly, various key scene optimization technologies are explored, including different scene data representation modes, mobile scene representation optimization, and mobile scene adaptive scheduling. By implementing these technologies, smartphones with different levels of performance can achieve higher frame rates for rendering scenes and improved image quality. Finally, a plugin-free prototype system is developed using a flood scenario for experimentation. The results show that the proposed methods achieve high enough scene rendering frame rates to meet the requirements for rendering 3D disaster scenes in mobile virtual reality [13].

The rapid growth of data volumes poses significant challenges in various applications such as medical imaging, physical modeling and industrial scanning. This growth can be attributed to the advancement of high-resolution scanners and the availability of high-performance graphics processing units (GPUs) that enable interactive visualization of large data sets. However, the increase in problem sizes is outpacing the growth of on-chip GPU memory, and the resolution of traditional display systems has not kept pace with the exponential growth in processing power. To address these challenges, Mariam Bahameish [14] has developed an integrated approach that focuses on the efficiency of data representation using lattice-based methods and enhances the visualization capabilities of exploring such data. Facilities called Immersive Cabin and Reality Deck were built, along with a range of visualization techniques to address the challenges posed by the growing volume of data. Specifically, a computational fluid dynamics (CFD) simulation framework based on the lattice Boltzmann method was developed using a face-centered cubic (FCC) lattice for stable simulations with optimal sampling efficiency. The simulation code is integrated into a visualization environment that includes a high-performance volumetric renderer and support for virtual reality systems. Volume rendering is further improved with a new LOD scheme that allows efficient mixing of optimal sampling lattices at different levels of the hierarchy. The author also developed the Reality Deck, a 1.5 gigapixel immersive display, and extended the visualization framework to support it. In addition, rendering techniques have been developed, such as conformal rendering for partially occluded virtual reality environments and frameless rendering, which replaces traditional frame buffers with reconstructed samples. Infinite Canvas technology allows you to explore gigapixel datasets while seamlessly updating graphics beyond the user's field of view. These methods are integrated into the Immersive Cabin and Reality Deck visualization platforms, providing advanced visualization capabilities for large and complex data [14].

Tiled video wall systems have generated significant interest in visualizing massive data sets from Butcher and Ritsos [15], offering an immersive and collaborative environment with high-resolution capabilities. To achieve efficient rendering, the rendering process must be parallelized and distributed across multiple nodes. The Data Observatory at Imperial College London has a unique setup with 64 screens running on 32 machines, delivering over 130 megapixels of resolution. Various applications, including ParaView, have been developed to use parallel rendering techniques and distributed rendering environments to achieve high-performance rendering. This project aims to leverage the potential of Data Observatory and ParaView in terms of visualization, advancing data exploration, analysis and collaboration through a scalable and high-performance approach. The basic concept involves setting up ParaView on a distributed cluster network and assigning the appropriate view to each screen by controlling the ParaView virtual camera. Application interaction events are broadcast to all connected nodes in the cluster to update their views accordingly. However, implementing such a system poses significant challenges, including synchronizing rendering across all screens, maintaining data consistency, and managing data partitioning [15].

Hanel et al. present their initial work developing prototypes of immersive analytics applications using new web technologies for virtual reality. They create 3D histograms that aim to resemble physical visualizations in the visualization community. The authors address the challenges faced by developers working with new virtual reality tools for the web and highlight the importance of creating effective and informative immersive 3D visualizations [16].

Kwon et al. argue that immersive virtual environments (IVEs) offer a suitable platform for visualizing and exploring 3D data, especially through the spatial understanding facilitated by stereo technology. However, compared to desktop setups, achieving lower latency and higher frame rates is crucial. The authors argue that existing implementations of direct volume rendering do not meet the desired rendering standards in terms of latency and visual quality without compromising immersion in the virtual environment. They review published acceleration methods and discuss their potential in IVE, focusing on head tracking as a key challenge and exploring optimization techniques. Traditionally, information visualization has been based on a 2D representation due to the dominance of 2D displays and reporting formats. However, the recent proliferation of consumer 3D displays and immersive head-mounted displays (HMDs) has opened new possibilities for immersive stereoscopic visualization environments. While immersive environments have been widely researched for spatial and scientific visualization, research in the context of information visualization has been limited. The authors present their thoughts on layout, rendering, and interaction techniques for visualizing graphs in immersive environments. They conducted a user study to compare their methods with traditional 2D graph visualization. The results show that participants using their methods answered questions significantly faster with fewer interactions, especially for more complex problems. Although overall correctness rates did not show significant differences, participants using their methods obtained significantly more correct answers for large graphs [17].

6. Proposed solution and implementation

One possible solution for distributed and parallel processing with rendering optimization for immersive data visualization in virtual reality is the use of client-server architecture in combination with optimized rendering techniques.

The following Table 2 identifies areas of optimization that add up to solutions to existing processing and rendering problems in a virtual environment.

Table 2
Description of optimization areas

№	Task	Description
1	Data distribution	A mechanism is required for dividing a large data set into smaller fragments and distributing them across different computing resources.
2	Parallel Processing	A mechanism for parallel/competitive processing of distributed data is required, for example, MapReduce, Apache Spark.
3	Rendering optimization	Leverage existing rendering optimization solutions such as Level of Detail (LOD), occlusion clipping, and adaptive rendering.
4	Load Balancing	A mechanism is required to balance the load between different computing nodes.
5	Caching and Preprocessing	A mechanism is required for data caching, as well as data preprocessing. For example, you can use Redis for caching.
6	Network load	A mechanism for effective data communication between the final visualization device and data processing servers/nodes is required.

6.1. Description of the proposed solution

Figure 2 shows the architecture of the proposed solution.

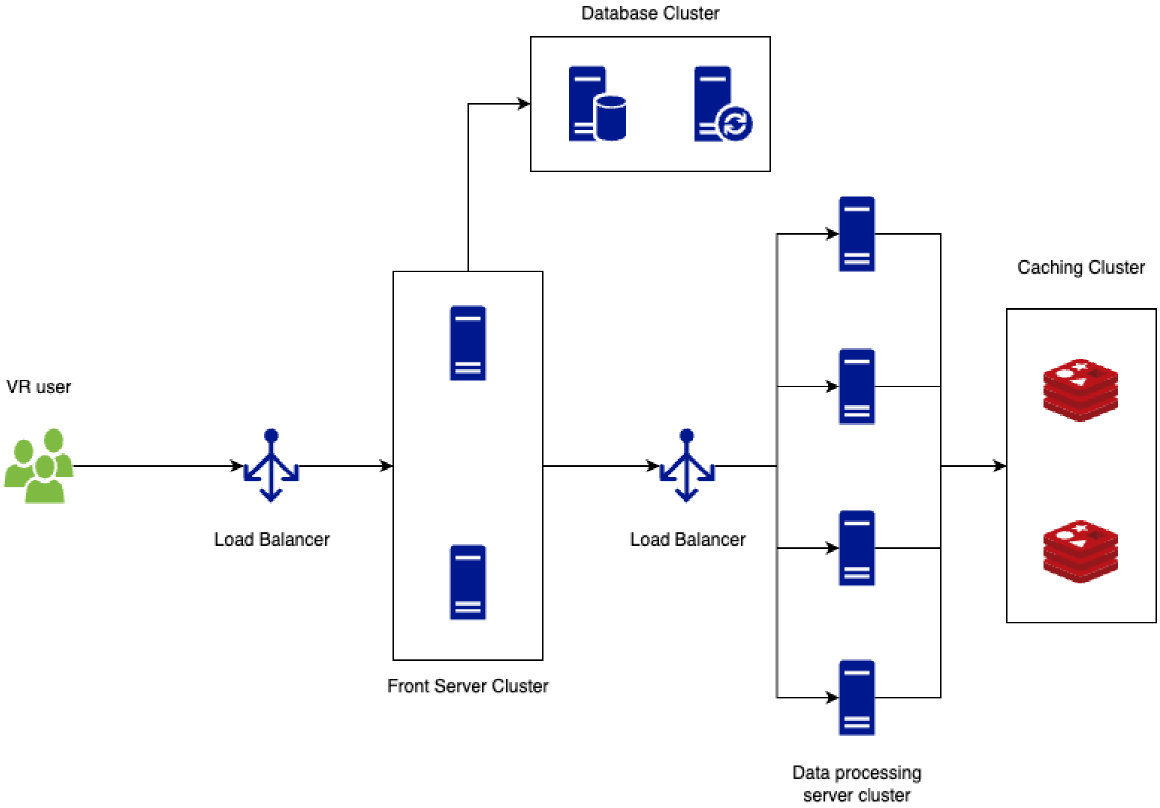


Figure 2: Architecture of the proposed solution

In this client-server architecture, a group of users “VR user” access data through a load balancer to the “Front Server Cluster”. While the data server reads data from the “Database Cluster” databases. After which the data is transferred to the “Data processing server cluster”. In this cluster, each node processes its own portion of data. The load on these nodes is also balanced. During processing, the data is necessarily cached, i.e. recorded in the “Caching Cluster”.

To maintain a high level of availability and reliability, all services are clustered and used in a High-Load configuration.

Table 3 provides descriptions of the functional responsibilities of the architecture components.

Table 3
Functional responsibilities of architecture components

№	Component	Task	Description
1	Front server	Data separation	1. Dividing a large data set into smaller pieces based on a predefined partitioning strategy. 2. Assigning each fragment to a computing node in a distributed system.
		Network communication	Optimized protocols for interaction with the end client.
2	Data processing server cluster	Parallel Processing	1. Each server processes its assigned pieces of data independently using parallel processing techniques.

			2. Applying the desired data mining and computational algorithms to extract relevant information from the data.
		Cleaning	After data processing, it is mandatory to release resources.
3	Load Balancer	Load Balancing	1. Load balancing mechanisms to evenly distribute the computing load between computing nodes. 2. Dynamically assign new pieces of data as needed.
4	Caching server	Data Caching	Caching preliminary and intermediate results, as well as frequently used objects.
5	VR user software	Rendering	Optimization of rendering of incoming data.
		Network communication	Optimized protocols for interaction with the end client.

7. Conclusion

There is a need to model and develop virtual environments for univariate and multivariate data analysis based on immersive visualization. The lack of software for effective data analysis through visualization increases the needs of both scientific and applied fields, including business, medicine, industry, logistics, etc. The relevance of the project is reinforced by the fact that the ongoing research expands the capabilities of existing fields of science: Data Science, Big Data, Data Analytics, Virtual Reality/Augmented Reality/Extended Reality. This article outlined the differences between traditional and immersive visualization, their capabilities, and limitations, as well as in-depth information on the issues that arise with immersive data visualization. The article provides a possible solution to existing problems, including architectural, conceptual, and technical approaches to the development of such systems.

8. References

[1] Srivastava, D. (2023). An Introduction to Data Visualization Tools and Techniques in Various Domains. International Journal of Computer Trends and Technology 71(4). pp. 125-130. doi:10.14445/22312803/IJCTT-V71I4P116.

[2] Kachan, H. (2022). Business analysis and data visualization means. Pryazovskyi Economic Herald. doi:10.32840/2522-4263/2022-1-11.

[3] Wedpathak, G. S., Nassa, V. K. (2023). Visualization Techniques - A Comparative Study in Data Science Workflows. Conference: INTERNATIONAL CONFERENCE ON 6G AND WIRELESS NETWORK TECHNOLOGIES.

[4] Senkuvienė, I., Jankauskas, K., Kvietkauskas, H. (2014). Using manufacturing measurement visualization to improve performance. Mechanika 20(1). doi:10.5755/j01.mech.20.1.3514.

[5] Caban, J. J., Gotz, D. (2015). Visual analytics in healthcare – opportunities and research challenges. Journal of the American Medical Informatics Association 22. pp. 260-262. doi:10.1093/jamia/ocv006.

[6] Comerón, T. M., Isabel, A. B., Sánchez, A. G. (2022). Visualization of information and knowledge to communicate social sciences research results.

[7] Huang, Z., Zheng, Y., Xie, L., Yang, T. (2006). A Virtual Plant Ecosystem Featuring Parallel Computing and Distributed Visualization. IEEE Xplore Conference Paper. doi:10.1109/IMSCCS.2006.25.

[8] Chen, H., Chen, Y., Finkelstein, A., Funkhouser, T., Li, K., Liu, Z., Samanta, R., Wallace, G. (2001). Data distribution strategies for high-resolution displays. Computers & Graphics 25. 811-818. doi:10.1016/S0097-8493(01)00123-6.
[9] Varga, M. N. (2023). Immersive Multidimensional Data Visualization using Geon Based Objects. University of Plymouth Research Theses, Research Theses Main Collection. doi:10026.1/20194.
[10] Flatken, M., Schneegans, S., Fellegara, R., Gerndt, A. (2023). Immersive and Interactive 3D Visualization of Large-Scale Geo-Scientific Data. IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 211-215. doi:10.1109/VRW58643.2023.00052.
[11] Cedilnik, A., Geveci, B., Moreland, K., Ahrens, J., Favre, J. (2006). Remote Large Data Visualization in the ParaView Framework. Eurographics Symposium on Parallel Graphics and Visualization.
[12] Hu, Y., Zhu, J., Li, W., Zhang, Y., Zhu, Q., Qi, H., Zhang, H., Cao, Z., Yang, W., Zhang, P. (2018). Construction and Optimization of Three-Dimensional Disaster Scenes within Mobile Virtual Reality. ISPRS International Journal of Geo-Information. 7. 215. doi:10.3390/ijgi7060215.
[13] Petkov, K. (2013). Lattice-based Immersive Visualization.
[14] Bahameish, Mariam. (2018). Scientific Data Visualization in an Immersive and Collaborative Environment. doi:10.5339/qfarc.2018.ICTPD147.
[15] Butcher, P. & Ritsos, P. (2017). Building Immersive Data Visualizations for the Web. Conference: 2017 International Conference on Cyberworlds (CW). 142-145. doi:10.1109/CW.2017.11.
[16] Claudia, H., Benjamin, W., Bernd, H., Torsten W. K. Interactive Volume Rendering for Immersive Virtual Environments. JARA – High Performance Computing.
[17] Kwon, O., Muelder, C., Lee, K., Ma, K. A Study of Layout, Rendering, and Interaction Methods for Immersive Graph Visualization. IEEE Transactions on Visualization and Computer Graphics 22(7). 1802-1815. doi:10.1109/TVCG.2016.2520921.