Assessing the Applicability of Virtual Reality for Data Visualization

Dissertation zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Matthias Kraus

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Informatik und Informationswissenschaft

Konstanz, 2021

Tag der mündlichen Prüfung: 25. November 2021

1. Referent: Prof. Dr. Daniel A. Keim

2. Referent: Prof. Dr. Michael Sedlmair

3. Referent: Prof. Dr. Harald Reiterer

Abstract

Data visualization is a powerful tool to efficiently and effectively extract knowledge from data and gain insights through interactive analysis procedures. For each data and task, there are different visualization techniques that are optimal for the particular analysis procedure. However, not only the visualization can affect the task - but also how the visualization is observed. It makes a huge difference if the data is analyzed on the small displays of smartphones, on standard monitor screens, on huge powerwall displays comparable to the size of cinema screens, or – what is a more recent development – in so-called immersive environments. In immersive environments, the real environment is augmented or replaced by virtual elements, allowing a high degree of freedom in the design of analysis environments, visualizations, and interaction concepts. Immersive analytics comprises analysis procedures that make use of such immersive and more engaging media. Currently, one of the biggest problems in immersive analytics is that the added value of immersive devices, for example, compared to conventional monitor screens, is not clearly defined. There is no common set of rules to determine, for instance, when it makes sense to use virtual reality. This is because the technology is relatively new and constantly evolving. Therefore, a regularly recurring challenge is to evaluate such immersive environments for information visualization applications and to answer the question of whether and how the deployment of virtual reality makes sense for a given combination of data, task, and visualization. This thesis describes how the assessment of added value through virtual reality can be carried out using three different strategies. The first is based on logical reasoning using deductions from literature to justify the use of immersive media without an evaluation in the classical sense. As a side effect, this approach can reveal research gaps and indicate which research directions are promising and should be further investigated. The second strategy examines a single property that is affected by immersive environments. As exemplary evaluations, two human factors, namely immersion and orientation, are considered and assessed in quantitative user studies. The third strategy involves holistic evaluations of immersive analytics applications. The implementation of this strategy is illustrated by means of a qualitative and a quantitative evaluation of two applications. Finally, all three strategies are discussed in terms of the sustainability and generalizability of their results. To date, immersive analytics has been subject to relatively little research, in part because it is a rapidly evolving field. The research presented in this dissertation is intended to help fill various research gaps by highlighting different approaches to where and how virtual reality can and cannot be put to good use.

Zusammenfassung

Visualisierung ist ein mächtiges Hilfsmittel, um in interaktiven, visuellen Analysen effizient und effektiv neue Erkenntnisse aus großen Datenmengen zu gewinnen. Es gibt eine Vielzahl unterschiedlicher Visualisierungstechniken, und durch die Ergebnisse jahrzehntelanger Forschung sind gewisse Leitlinien entstanden, die konkretisieren, welche Techniken für eine bestimmte Kombination aus Daten und Analysezielen erfolgversprechend sind. Doch nicht nur die Visualisierungstechnik kann Einfluss auf eine Analyseaufgabe haben, sondern auch die Art und Weise, wie die Visualisierung betrachtet wird. Es macht einen großen Unterschied, ob die Daten auf dem kleinen Display eines Handys, einem Monitorbildschirm, einer Powerwall, oder - was eine neuere Entwicklung darstellt - in sogenannten immersiven Umgebungen analysiert werden. In immersiven Umgebungen wird die reale Umgebung durch virtuelle Elemente erweitert oder ersetzt, was zu einem hohen Freiheitsgrad bei der Gestaltung von Analyseumgebungen, Visualisierungen und Interaktionskonzepten führt. Analyseverfahren, die sich solcher immersiver Medien bedienen, können unter dem Begriff 'Immersive Analytics' zusammengefasst werden. Eines der größten Probleme bei Immersive-Analytics-Ansätzen besteht derzeit darin, dass der Mehrwert des Einsatzes eines immersiven Mediums, beispielsweise gegenüber einem herkömmlichen Bildschirm, nicht klar definiert ist. Es gibt keine allgemeingültigen Richtlinien, um zum Beispiel zu bestimmen, wann der Einsatz von Virtual Reality sinnvoll ist. Der Grund hierfür ist, dass die Technologie relativ neu ist und sich ständig weiterentwickelt. Eine Herausforderung besteht daher darin, immersive Umgebungen für Informationsvisualisierungsanwendungen kontinuierlich zu evaluieren und die Frage zu beantworten, ob und auf welche Weise es Sinn macht, Virtual Reality für eine bestimmte Kombination von Daten, Analyseaufgaben und Visualisierungen einzusetzen. In dieser Dissertation werden drei Strategien vorgestellt, mit deren Hilfe die Anwendbarkeit von Virtual Reality für eine bestimmte Analyse beurteilt werden kann. Die erste basiert auf logischer Argumentation mittels Ableitungen aus bestehender Literatur, um den Einsatz immersiver Medien ohne eine Evaluierung im klassischen Sinne zu rechtfertigen. Als Nebeneffekt kann dieser Ansatz Forschungslücken aufdecken und aufzeigen, welche Forschungsrichtungen vielversprechend sind und weiter untersucht werden sollten. Die zweite Strategie untersucht eine einzelne Eigenschaft, die durch immersive Umgebungen beeinflusst wird. Als beispielhafte Evaluationen werden zwei menschliche Faktoren, Immersion und Orientierung, in quantitativen Nutzerstudien betrachtet und bewertet. Die dritte Strategie konzentriert sich auf die ganzheitliche Bewertung einer Immersive-Analytics-Anwendung. Die Umsetzung dieser Strategie wird anhand einer qualitativen und einer quantitativen Evaluation von zwei Anwendungen veranschaulicht. Alle drei Strategien werden im Hinblick auf die Nachhaltigkeit und Verallgemeinerbarkeit ihrer Ergebnisse diskutiert. Bislang wurde Immersive Anlytics relativ spärlich erforscht, unter anderem weil es sich hierbei um ein sich schnell entwickelndes Forschungsfeld handelt. Die in dieser Dissertation vorgestellte Forschung soll dazu beitragen, einige Forschungslücken zu schließen, indem verschiedene Ansätze aufgezeigt werden, wo und wie Virtual Reality sinnvoll eingesetzt werden kann und wo nicht.

Acknowledgements

I want to thank everyone that supported me over the last years. In particular, I am grateful for my two mentors Daniel Keim and Michael Sedlmair: Thank you, Daniel, for letting me join your workgroup, laying the foundation for my ambitions, and always having an open door for feedback and advice; Also thanks to you, Michael, for the many research collaborations, all the advice and guiding, and the continuous stream of constructive feedback. Thanks to all members of our research group in Konstanz who have made my time as a doctoral student entertaining and more interesting due to great collaborations, exciting conversations, and sometimes even Skat tournaments. Even though in the last two years, the social aspect came off short due to the Corona pandemic, we have made good advance performance in the years before – team building events, Christmas parties, summer BBQs, and retreats were always great! Thank you, Johannes Fuchs, Juri Buchmüller, Mennatallah El-Assady, Wolfgang Jentner, Matthias Miller, Hanna Schäfer, Fabian Sperrle, Thilo Spinner, Michael Blumenschein, Maximilian Fischer, Florian Stoffel, Dominik Sacha, Michael Behrisch, Philipp Meschenmoser, Evanthia Dimara, Alexandra Diehl, Niklas Weiler, Thorsten Breitkreutz, Johannes Häußler Dirk Streeb, Udo Schlegel, Bruno Schneider, Daniel Seebacher, Rita Sevastjanova, Michael Behrisch, Eren Cakmak, Frederik Dennig, Yannick Metz, Thomas Polk, Manuel Stein, Sabine Kuhr, Matt Sharinghousen, Benedikt Bäumle, Timon Kilian, and Daniel Schweitzer. Further, thanks to all my co-authors outside of the group - I enjoyed the exciting conversations and fruitful collaborations with you! Thank you, Leonel Merino, Karsten Klein, Falk Schreiber, Harald Reiterer, Min Chen, Thomas Pollok, Tobias Moritz, Chengchao Qu, Katrin Angerbauer, Tobias Schreck, Jürgen Bernard, Björn Sommer, Ulrich Engelke, Daniela Oelke, Johannes Kehrer, Benjamin Bach, Magdalena Schwarzl, Dieter Schmalstieg, Daniel Weiskopf, and David Schubring. I also would like to thank my family and friends who always had my back and provided very welcomed distractions from work. In particular, thanks to you, mum, dad, Alex, and Eli! Last but not least, special thanks to my beautiful wife Tanja, who always supported me (mentally and practically as in proofreading) and was the source of positive energy that made this work possible. Thank you for being part of my life!

Table of Contents

1 Introduction

Contents

In this first chapter, the research objectives are motivated and related to open research problems and areas. The introduction also serves to define and explain terms and concepts used throughout the thesis and to illustrate the structure of the following chapters. Additionally, it comprises a list of all publications used in this thesis, clarifying the division of work among all co-authors of each paper.

1.1 Motivation & Research Objectives

Immersive analytics (IA), which refers to research on analyses concerned with the "use of engaging, embodied analysis tools to support data understanding and decision making" [102], has repeatedly experienced an upswing in research interest over the past few decades. What peaked in the late 90s with immersive CAVE environments reached another peak in the last decade (late 2010s) with head-mounted displays (HMDs). This shows that the field is very diverse and is not tied to a specific technology - but rather depends on specific properties of the respective technology. Over the years, more and more niches of applications have been found where the use of immersive environments can be advantageous, for example, tasks that strongly benefit from improved distance and structure perception, better spatial understanding and orientation, or even the minimization of distractions from the real environment. Properties inherent in new technologies, such as direct interaction, stereoscopic vision, and immersion, have proven beneficial for some visual analysis tasks. However, immersive analytics is by no means the panacea that will supersede classical methods of visual analysis and can be deployed for every problem.

Previous work has shown that the deployment of immersive devices can be associated with beneficial factors, which in turn could mitigate some drawbacks and increase the efficiency and effectiveness of certain visualization approaches. However, the cost-benefit ratio must be considered, and not every flaw is automatically cured. For example, 3D visualizations that have proven to work poorly on screens, such as 3D pie charts, do not suddenly become great visualization tools just because they are inspected through virtual reality (VR) HMDs. At the same time, the use of immersive technologies does not mean that the visualizations presented have to be 3D or that classic interaction modalities (e.g., keyboard and mouse) have to be abandoned. The most prominent preconception regarding immersive analytics is that visualizations must be displayed in 3D. Due to the fact that the visualization design space is 3D, we actually have to make use of all three dimensions for our visualizations, and therefore immersive analytics is synonymous with 3D visualization. However, immersive analytics is much more than just deploying 3D visualizations on a stereoscopic display. It is about finding application areas where the application of immersive technologies adds value compared to a conventional visualization or where certain induced disadvantages are outweighed by simultaneously induced advantages.

Therefore, it is important to find ways to assess the applicability of VR for a specific use case. That is, a strategy to estimate whether the deployment of VR adds value or not. The focus of this work is on such assessment strategies that are at a higher meta-level compared to evaluation approaches. The goal of assessment is to arrive at a reasoned decision, whereas the goal of evaluation is to arrive at a grounded theory. Cohen et al. [68] describe grounded theory as the result of an evaluation process that verifies a hypothesis based on empirical evidence. Optimally, the assessment is based on grounded theory, but it can also be based on less than empirical evidence, such as logical reasoning or deductions from similar scenarios. In the field of data visualization alone, many efforts have been made to analyze, structure, classify, review, and quantify evaluation approaches. To name a few examples: Carpendale [52] discusses challenges in the empirical validation of data visualization, divides the evaluation spectrum into different types of evaluation, and discusses the advantages and disadvantages of different empirical methodologies. Isenberg et al. [171] present a systematic review of the practice of evaluations in the visualization domain and reflect, among other things, on issues that hinder cross-validations and reduce the validity of results. Lam et al. [216] reflect on seven different visualization evaluation approaches and discuss their benefits and drawbacks.

The ultimate goal would be to provide a universal guide that advises when immersive environments should and should not be used after evaluating all possible approaches. In contrast to the technological development of monitor screens, the technological progress of immersive devices is faster and makes much larger leaps that significantly impact the experience of the immersive environment. Especially due to the rapidly evolving field of immersive technologies, it is difficult to establish clear guidelines on where and how their application actually leads to benefits. For instance, the results of a study using a specific CAVE environment may not be applicable to a current VR HMD. Therefore, findings from evaluations of different immersive environments should be applied with caution to situations with a different technological basis. At best, they should be used as a starting point for formulating new hypotheses about one's own environment. Although research efforts have increased over the past decade to find answers to where and how the deployment of immersive technologies makes sense, the field is still largely unexplored.

The purpose of this thesis is to investigate how the added value of virtual reality for data visualization solutions can be assessed, taking into account the induced disadvantages and limitations of the deployed immersive environment. Therefore the driving question of this thesis is:

"How can the applicability of virtual reality for data visualization be assessed?"

We identified three ways to assess the applicability of VR and accordingly formulated the research objectives (O1-O3) of this thesis. Each is concerned with answering the *"How?"* in a specific way:

O1: through logical reasoning and the collection of arguments from literature

"Which arguments speak for or against the use of VR for data visualization?"

In this assessment strategy, opinions are formed through intensive literature research and logical reasoning. The result is a set of unconfirmed but well-founded hypotheses which can be used as a starting point for quantitative evaluations. This is the weakest form of assessment, as at best statements like the following can be formed: "VR should be advantageous in collaborative analysis tasks with realistic avatars, since it improves communication, as collaborators can see each other, interact in the same (virtual) space, and use gestures and facial expressions". Furthermore, the argumentation may contain limitations that go beyond the current state of the art and are thus speculative and hypothetical in nature.

O2: through low-level, single-aspect evaluation

"What distinguishes VR from conventional media and what effects does this have?"

The second way to assess the applicability of VR to a specific information visualization task is based more on empirical evidence. First, differences between conventional media and VR are investigated, searching for unique features of immersive environments that potentially influence omnipresent characteristics such as human factors. Subsequently, such a unique feature is examined in more detail by evaluating its impact on a particular characteristic in a particular immersive analytics task. For example, the feature could be 'natural body movements' in VR, the investigated characteristic 'orientation' in the virtual space, and the task 'wayfinding' in the virtual environment. Typically, this is done in direct comparisons between screen and VR setups. The goal should be to isolate individual factors and to design the study conditions on the screen as similar as possible to the conditions in VR. In this form of assessment, a specific aspect is evaluated in a specific context. When attempting to transfer findings from such evaluations to another application, one must be aware that the altered context could have a significant impact on the desired effects. For example, when studying orientation skills, a maze environment could be used to compare the orientation of users while playing a game on a screen or in VR. The result could be that orientation is better in VR due to spatial memory. However, if one wants to exploit this benefit for another task, e.g., in an immersive scatterplot visualization, the offset in orientation between screen and VR might not be large or even nonexistent due to the new environment. In summary, it is difficult to determine all influencing factors that lead to a certain result and thus to generalize findings.

O3: trough high-level, big-picture application assessment

"How can we exploit properties of VR to generate benefits for visual analytics applications?"

The third assessment strategy is the high-level evaluation of a particular application. While it may be possible to compare individual aspects of the application when perceived through different media, it is often not feasible to compare the overall application in a direct lineup. Even for techniques and applications where a comparison with an analog, screen-based application or technique is possible, a quantitative comparison is questionable because many factors are changed, and the "analog" system could be completely different in design. Therefore, this assessment strategy mainly focuses on a qualitative assessment of the immersive environment for a certain task and can indicate where the application of VR leads to promising results. In contrast to low-level property evaluations, this approach has the advantage of directly assessing the applicability of VR to a particular information visualization task, but the disadvantage that it is not very controllable whether revealed benefits can be directly attributed to immersive environments.

In summary, there are three different strategies with different strengths and weaknesses how to assess the applicability of VR environments. In this thesis, each strategy is demonstrated with several examples. But before we get to these examples, we clarify the terminology used, give a more detailed overview of the structure, and provide a list of all publications that are presented in parts in this thesis.

1.2 Terminology & Definitions

In this section, several key terms and concepts used throughout the thesis are deduced from related literature and summarized. For many terms, such as VR, immersion, or presence, numerous definitions exist. The literature review aims at eliminating ambiguities by defining the terminology and thus paving the way for later lines of argumentation.

Virtual Reality (VR) and Virtual Reality Environments (VREs)

According to the dictionary, we can define the terms 'virtual' and 'reality' as follows: Virtual is something that is "temporarily simulated or extended by computer software" (e.g., virtual memory on a hard disk) [94]; reality is "something that exists independently of ideas concerning it" [93]. To compare VR, as used in computer science, with conventional media, we must first specify what exactly is meant by a VR system in this context. The first experiments with VR technology were conducted several decades ago [119]. Since then, a vast amount of different VR equipment has been developed, ranging from pocket-sized cardboards like the Google Cardboard [136] to room-sized installations like the CAVE [76]. Different properties of VR prototypes may be the reason why several definitions have been introduced in recent years. Although most of them are relatively similar, they differ in one crucial aspect: They describe the defining properties of VR from different perspectives. We could identify three general types of definitions for virtual reality.

Hardware-centered VR Some VR definitions focus mainly on the hardware aspect. They usually include some kind of stereoscopic display as well as interaction controllers or data gloves. A representative definition of this category was introduced by Stephen Ellis [104], who links virtual reality to the hardware used to create it – a head-mounted display. VR displays track users' head positions and adapt themselves accordingly, enabling users to navigate through the virtual reality environment. VR is created by VR hardware.

Human-centered VR Latta and Oberg [218] provide an example of what we call a human-centered definition of virtual reality. They define virtual reality as an interplay between hardware and user. The hardware monitors human behavior and, in turn, stimulates the human perceptual and muscle systems. Their model of a VR system consists of effectors and sensors. Effectors describe the hardware that the VR system uses to stimulate the human body, while sensors are responsible for detecting human actions. VR is created through the synergy of human and computer.

Concept-centered VR Concept-based approaches deduce virtual reality from a conceptual construct. For instance, Steuer [359] derives VR from the concept of telepresence. According to Steuer, someone is in a virtual reality if he or she is telepresent in that environment. He defines telepresence as presence originating from a communication medium (e.g., a VR headset). Furthermore, the degree of telepresence (and thus immersion) depends on two dimensions: vividness and interactivity. Vividness describes how well the virtual world emulates the real world, and interactivity describes how much the user can influence it. Virtual reality is created by applying a concept to the user (e.g., telepresence: mentally transporting users to a different place).

Immersion and Presence

Each of these VR definitions addresses, explicitly or implicitly, the two key aspects of VR: immersion and presence. They describe the states of a user located in a virtual reality environment. Even though both terms are key elements of VREs, there is an ongoing discussion about their meaning and definition. For example, Cummings and Bailenson summarize the concept of presence as the feeling of "being there" in a mediated virtual environment. Presence is responsible for the effectiveness of VR applications as it magnifies stimuli and virtual interactions (human-centered) [78]. McGloin et al. state that the term 'immersion' is synonymous with presence [241] and offer the same definition as Cummings and Bailenson.

According to Slater and Wilbur, the term 'immersion' describes the technological basis used to implement a VRE that influences its medial quality (hardware-centered) [348]. Immersion comprises the ability of a system to provide persuasive conditions that are perceived by human sensory organs. Consequently, the degree of 'immersion' of a system is an objective measure that can be used to compare the quality of different systems. Immersion depends on the ability of a system to fade out the physical reality by appealing to all senses of an immersed person as much as possible. The richness of provided information ultimately affects the vividness of the VRE (concept-centered) as a function of screen resolution and fidelity. Besides, self-perception is another crucial aspect of immersion since the presence of a virtual body in a VRE creates the perception of being part of the constituted environment.

Witmer et al. describe presence as the subjective experience of being transported to another place or environment (concept-centered)[413]. The intensity of presence influences the sensation of an individual experiencing the VRE instead of the actual environment. Ijsselsteijn and Riva consider the experienced presence to be a complex and multidimensional perception shaped by various cognitive processes and the composition of sensory information (human-centered) [41]. The perceived degree of presence depends on the immersed person's ability to have control over the mediated information. The simultaneous stimulation of multiple sensors affects cognition, perception, and emotion, which are key components of presence. The better these dimensions are addressed by the VR system, the higher the degree of presence that can be experienced in a VRE.

According to this extensive amount of recent work, optimizing presence is the primary objective when designing a VRE. Since presence directly depends on the degree of immersion, technological improvements that enhance perceived immersion (e.g., an increase in resolution) also augment presence [31]. High degrees of immersion provide users with an improved psychological experience of "being there" [78]. The sensation of presence in a VRE can vary greatly from person to person, as it depends on individual factors such as attentional resources and physical conditions.

1.3 Thesis Structure & Scientific Contributions

As shown in Figure 1.1, the thesis is organized into six chapters, three of which are main sections (3, 4, 5). The main sections are enclosed by a background chapter (2) at the beginning, which provides an overview of previous research in the field and current research gaps, and a concluding chapter (6) at the end, which summarizes the research presented. In the following, the scientific contributions (*C1-C10*) of this thesis are outlined along this structure.

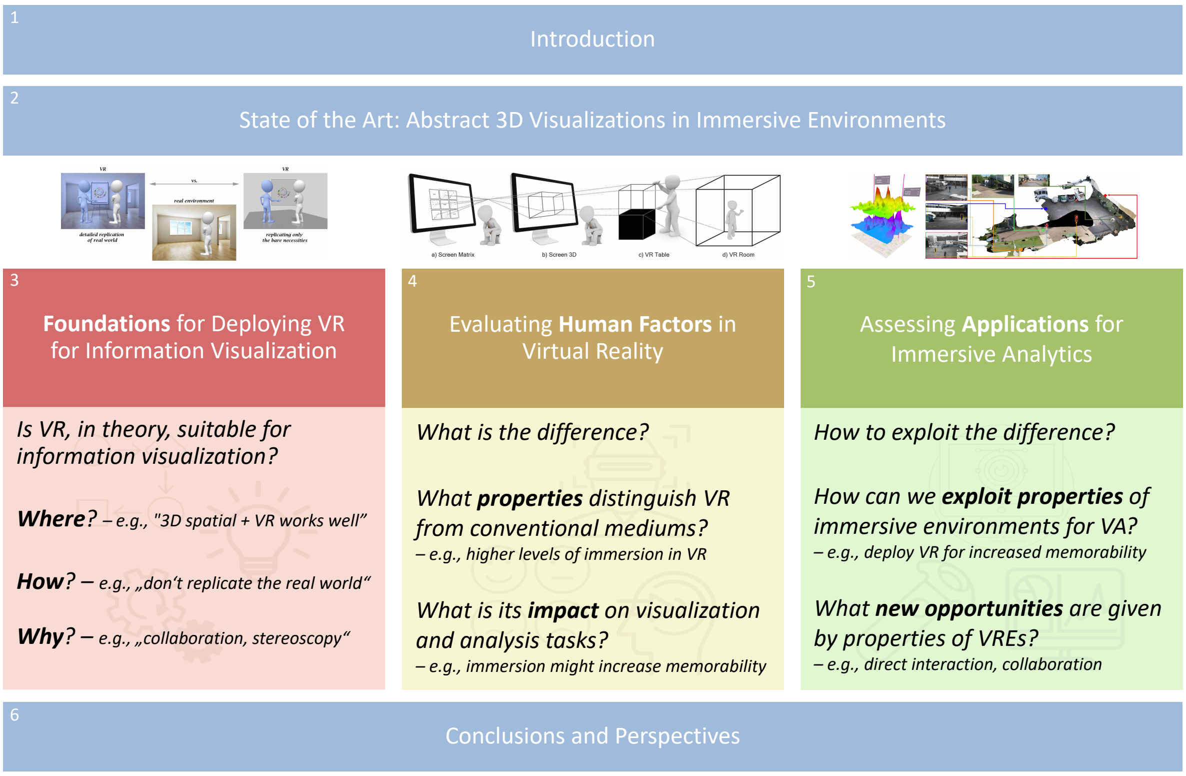

Figure 1.1: The thesis is structured into six chapters.

***Chapter 1* (Introduction)**

. . . introduces and motivates the research presented in this thesis.
The first chapter provides a rationale for the research conducted, clarifies the terminology used, introduces the structure, and points out all parts of this thesis that have been previously published.

***Chapter 2* (State of the Art: Abstract 3D Visualizations in Immersive Environments)**

. . . provides the big picture by giving an overview of the field and reveals research gaps.
The second chapter provides an overview of the research context by means of a literature survey of previous work addressing the use of immersive environments for abstract 3D visualizations, which is the most controversial application area *(C1)*. At the same time, the chapter identifies open research areas and promising research directions and discusses implications of findings from previous work on information visualization *(C2)*. The resulting landscape of research gaps serves as a starting point for subsequent chapters. A detailed context is provided in each chapter individually since the research directions covered are quite diverse.

Chapter 3 **(Foundations for Deploying Virtual Reality for Data Visualization)**

... **provides the argumentative basis for further research in this area and reveals versatile application areas and opportunities of VR.**

The first main chapter considers the use of virtual reality for information visualization from a general argumentative point of view. First, the "curse of visual data analysis" *(C3)* is introduced as a potentially problematic discrepancy between the data world and the real world, and two strategies *(C4)* for overcoming this 'curse' by minimizing the visual offset between the two worlds are pointed out. Subsequently, the principle equivalence of virtual reality with conventional media is discussed and demonstrated by means of thought experiments and a clear line of argumentation *(C5)*. The chapter concludes with a discussion of design guidelines for immersive environments on the meaningfulness of replicating real-world properties and circumstances in the virtual environment *(C6)*.

Chapter 4 **(Evaluating Human Factors in Virtual Reality)**

... **demonstrates evaluation approaches by presenting two detailed evaluations of aspects inherent to or affected by immersive environments: immersion and orientation.**

This chapter presents results of a user study comparing four visualization design spaces that differ with regard to their level of immersion *(C7)*, as well as results of a user study comparing three visual orientation-supporting tools in terms of their suitability for supporting a user's orientation in a virtual environment *(C8)*.

Chapter 5 **(Assessing Applications for Immersive Analytics)**

... **demonstrates the applicability of VR for data visualization by presenting the assessment of two immersive analytics applications.**

The third main chapter introduces a new technique *(C9)* for the comparative analysis of 2D distributions (stacked 3D heatmaps) and presents its evaluation in the form of a user study comparing the new technique with a conventional juxtapositioned comparison approach. Additionally, this chapter presents a vivid application of VR for 4D scene reconstruction and exploration *(C10)*.

Chapter 6 **(Conclusions and Perspectives)**

... **summarizes the contributions of the research presented and reflects on perspectives.**

The final chapter completes the dissertation by reflecting on the research area on a meta-level and placing the previously presented assessments in the broader context of their research objectives. At the same time, the main contributions are summarized, and open research areas, as well as promising future research directions, are outlined.

1.4 Publications

Thesis-Relevant Publications and Contribution Clarification

Parts of this thesis have been published in conference proceedings and book chapters. The following list gives an overview of all publications contributing to this thesis and shows the distribution of the authors' contributions. Publications are arranged chronologically.

- [200] Matthias Kraus, Johannes Fuchs, Björn Sommer, Karsten Klein, Ulrich Engelke, Daniel A. Keim, Falk Schreiber (2021). **Immersive Analytics with Abstract 3D Visualizations: a Survey**. *Journal Article*. In *Computer Graphics Forum*. *(Chapter 2)*
  This paper is the result of many years of work, as the paper project started back in 2017. In 2018, I took the main responsibility for this publication. In order to ensure the up-to-dateness of the paper, in addition to writing on the survey, we included newly published research every year. After a change of strategy at the end of 2019, I completed the classification of papers autonomously. The first draft was done by all the authors as we distributed the work among us. However, I did a major revision of the entire document, unifying the writing style and improving the reading flow. I then revised the document several times, taking into account feedback from all co-authors. Therefore, the paper is reused in chapter 2 without quotation marks.

- [201] Matthias Kraus, Karsten Klein, Johannes Fuchs, Daniel A. Keim, Falk Schreiber, Michael Sedlmair (2021). **The Value of Immersive Analytics**. *Journal Article*. In *IEEE Computer Graphics and Applications*. *(Chapter 6)*
  The idea for this paper emerged from a discussion with Michael Sedlmair. The original idea was then refined and elaborated with all co-authors in brainstorming sessions. Using digital whiteboards, we organized our ideas and summarized them into four guiding scenarios, which are described in the paper. I wrote the first draft of the entire paper and refined it iteratively with feedback from all co-authors. Therefore, I reuse parts of the paper in this thesis without citations in section 6.1.

- [203] Matthias Kraus, Matthias Miller, Juri Buchmüller, Manuel Stein, Niklas Weiler, Daniel A. Keim, Mennatallah El-Assady (2020). **Breaking the Curse of Visual Analytics: Accommodating Virtual Reality in the Visualization Pipeline**. *Book Chapter*. In *Communications in Computer and Information Science (Springer)*. *(Chapter 3)*
  The book chapter is a follow-up to our conference paper on using visual context expansion to improve uncertainty analysis and result verification [206]. The focus of the extension is shifted to VR and the implications of moving from conventional media to VR. The main contributions of the paper were developed in several group discussions, mainly between M. Miller, J. Buchmüller, N. Weiler, M. El-Assady, and myself. Together with M. Miller, I conducted the qualitative user study on the effectiveness of one presented use case prototype (6.5) with law enforcement agents of the federal police headquarters (Bundespolizeipräsidium) in Potsdam. I wrote the first draft of the entire paper and refined it iteratively with feedback from all co-authors. Therefore, in this thesis I reuse the paper without citations in section 1.2 and section 3.2. Repetitive elements of the underlying model, already discussed in section 3.1, have been edited out.

- [205] Matthias Kraus, Hanna Schäfer, Philipp Meschenmoser, Daniel Schweitzer, Daniel A. Keim, Michael Sedlmair, Johannes Fuchs (2020). **A Comparative Study of Orientation Support Tools**

in Virtual Reality Environments with Virtual Teleportation. *Conference Proceedings*. In *Proceedings of the International Symposium on Mixed and Augmented Reality*. *(Chapter 4)*
I took the main responsibility for this publication and initiated the paper project to investigate the research question if, how, and when visual tools can promote orientation in virtual reality environments when virtual teleportation is deployed as a movement compensating technique. In several group sessions, H. Schäfer, M. Sedlmair, J. Fuchs, and I developed the study design, which was put into action by D. Schweitzer and me developing the corresponding study prototype and conducting the user study. I evaluated the study results and produced a first full draft. All co-authors provided feedback, which I implemented in several revisions. Therefore, the paper text is reused in section 4.2 without citation marks.

- [198] Matthias Kraus, Katrin Angerbauer, Juri Buchmüller, Daniel Schweitzer, Daniel A. Keim, Michael Sedlmair, Johannes Fuchs (2020). **Assessing 2D and 3D Heatmaps for Comparative Analysis: An Empirical Study**. *Conference Proceedings*. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. *(Chapter 5)*
I took the main responsibility for this publication. The original idea for the technique developed throughout several discussions with J. Buchmüller, D. Schweitzer, and me during the directed studies course “Applications for Powerwall and Virtual Reality Environments”. The idea was continuously improved through further discussions with all co-authors and put into practice. I developed a study design to compare the new technique with a conventional comparative analysis method under consultation with J. Fuchs, K. Angerbauer, M. Sedlmair, and J. Buchmüller. With the help of D. Schweitzer, we then developed a study prototype tailored to the study design. Together with K. Angerbauer, I conducted the study. I then evaluated the study results and produced a complete first draft. Taking into account feedback from all co-authors, I revised the paper several times. Therefore, the paper is reused in section 5.1 without citation marks.

- [204] Matthias Kraus, Thomas Pollok, Matthias Miller, Timon Kilian, Tobias Moritz, Daniel Schweitzer, Jürgen Beyerer, Daniel A. Keim, Chengchao Qu, Wolfgang Jentner (2020). **Toward Mass Video Data Analysis: Interactive and Immersive 4D Scene Reconstruction**. *Book Chapter*. In *Sensors; Special Issue Selected Papers from the 9th International Conference on Imaging for Crime Detection and Prevention (ICDP-19)*. *(Chapter 5)*
This book chapter is a major extension of a previous conference paper [289] on the research prototype developed within the EU project VICTORIA. The focus of the extension is shifted to the VR capabilities of the prototype and related application areas and possibilities. The described prototype was developed in close collaboration with IOSB Fraunhofer Institute (T. Pollok, T. Moritz, C. Qu). The complete concept and pipeline for 4D crime scene reconstruction from mass image and video data was developed by T. Pollok. All presented scene reconstructions were created by T. Pollok. The visual interface, visualization elements, and the integration of VR were developed by myself, M. Miller, W. Jentner, and T. Kilian. The first draft of the paper was prepared by several authors: W. Jentner (Introduction, 6.2), M. Miller (Related Work, 3.5), T. Pollok (3.2, 3.3), and myself (Abstract, Related Work, 3.1, 3.4, 3.6, 4, 5, 6.1, 6.3, 7). I incorporated feedback from T. Pollok, C. Qu, and W. Jentner and revised the entire paper several times. Therefore, with the exception of sections 2.1, 3.2, and 3.3., the text of the paper is reused in section 5.2 without citation marks.

- [206] Matthias Kraus, Niklas Weiler, Thorsten Breitkreutz, Daniel A. Keim, Manuel Stein (2019). **Breaking the Curse of Visual Data Exploration: Improving Analyses by Building Bridges between Data World and Real World**. *Conference Proceedings*. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (IVAPP).* *(Chapter 3)*
  The initial idea for the paper came from M. Stein, who initiated and supervised the paper project. The contributions of the paper resulted from many group discussions during the planning and writing phase. First drafts were written by M. Stein (Motivation, 4.2), T. Breitkreuz (Related Work, Use Case 1), Niklas Weiler (introduction to Section 4), and myself (remaining parts). The entire paper was revised iteratively several times by all co-authors. A final complete revision was done by me. Therefore, in this thesis I reuse the paper text without citation marks in section 3.1 and only removed the soccer example from the original section 4.1, which is from M Stein's work.

- [208] Matthias Kraus, Niklas Weiler, Daniela Oelke, Johannes Kehrer, Daniel A. Keim, Johannes Fuchs (2019). **The Impact of Immersion on Cluster Identification Tasks**. *Journal Article*. In *IEEE Transactions on Visualization and Computer Graphics.* *(Chapter 4)*
  This paper resulted from a project collaboration with Siemens AG (J. Kehrer, D. Oelke). Within the project we developed a suitable study design in several group discussions. Together with N. Weiler I developed the study prototype and conducted the study. I evaluated the study results and wrote a first draft together with N. Weiler. In order to extended the validity of the paper and to support the argumentative basis of the first study, I conducted a second study that exclusively examined differences between the previously considered design spaces of the first study. J. Kehrer, J. Fuchs, and D. Oelke provided useful feedback which was integrated in several revision circles. The final revision of the entire paper was done by me. Therefore, in this thesis I reuse the paper text without citation marks in section 4.1.

- [207] Matthias Kraus, Niklas Weiler, Alexandra Diehl, Benjamin Bach (2018). **Visualization in the VR-Canvas: How much Reality is Good for Immersive Analytics in Virtual Reality?** *Conference Proceedings*. In *Proceedings of VisGuides: 2nd Workshop on the Creation, Curation, Critique and Conditioning of Principles and Guidelines in Visualization at IEEE Vis.* *(Chapter 3)*
  The idea for this paper came from a forum discussion on VisGuides.org in which I discussed a guideline for designing virtual reality environments with several visualization researches. Extending this discussion, A. Diehl, B. Bach, N. Weiler, and I structured arguments, properties, and dimensions related to this discussion. The main contributions of this paper were developed in these group discussions. The first draft of the paper was written by several authors: B. Bach (Abstract, 3), A. Diehl (1, 4), and myself (2, 4, 5). The final paper was revised by myself and is, therefore, reused in section 3.3 without citation marks.

Additional Publications

The following list contains publications that I contributed to, but which were not included in this thesis.

- [246] Leonel Merino, Magdalena Schwarzl, Matthias Kraus, Michael Sedlmair, Dieter Schmalstieg, Daniel Weiskopf (2020). **Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009–2019)**. In *Proceedings of the 2020 ISMAR International Symposium on Mixed and Augmented Reality.*

- [333] David Schubring, Matthias Kraus, Christopher Stolz, Niklas Weiler, Daniel A. Keim, Harald T. Schupp (2020). **Virtual Reality Potentiates Emotion and Task Effects of Alpha/Beta Brain Oscillations**. In *Brain Sciences*.

- [289] Thomas Pollok, Matthias Kraus, Chengchao Qu, Matthias Miller, Tobias Moritz, Timon Kilian, Daniel Keim, Wolfgang Jentner (2019). **Computer Vision Meets Visual Analytics: Enabling 4D Crime Scene Investigation from Image and Video Data**. In *Proceedings of 9th International Conference on Imaging for Crime Detection and Prevention (ICDP-19)*.

- [251] Matthias Miller, Hanna Schäfer, Matthias Kraus, Marc Leman, Daniel A. Keim, Mennatallah El-Assady (2019). **Framing Visual Musicology through Methodology Transfer**. In *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS 2019*.

- [250] Matthias Miller, Johannes Häußler Matthias Kraus, Daniel A. Keim, Mennatallah El-Assady (2018). **Analyzing Visual Mappings of Traditional and Alternative Music Notation**. In *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS*.

- [321] Dominik Sacha, Matthias Kraus, Daniel A. Keim, Min Chen (2018). **Vis4ml: An Ontology for Visual Analytics Assisted Machine Learning**. In *IEEE Transactions on Visualization and Computer Graphics*.

- [321] Dominik Sacha, Matthias Kraus, Jürgen Bernard, Michael Behrisch, Tobias Schreck, Yuki Asano, Daniel A. Keim (2017). **SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance**. In *IEEE Transactions on Visualization and Computer Graphics*.

2

State of the Art: Abstract 3D Visualizations in Immersive Environments

Contents

After a long period of skepticism, more and more publications describe basic research as well as practical approaches on how to present abstract data in immersive environments for effective and efficient data understanding. Central aspects of this important issue in immersive analytics research are concerned with the use of 3D for visualization, the embedding in the immersive space, the combination with spatial data, suitable interaction paradigms, and the evaluation of use cases. In this chapter, the focus is on *abstract* data visualizations as they represent the most controversial application area in immersive analytics. The challenges and open research gaps discussed largely apply to non-abstract data visualization use cases as well and, thus, serve as a general basis for the following chapters. This chapter is based on our survey paper [200] which provides an overview of the state of the art in immersive analytics of abstract data.

2.1 Introduction

Immersive analytics is a field of research concerned with the design and application of engaging analysis tools to support data understanding and decision making [102]. It combines efforts from *scientific visualization (SciVis)*, *information visualization (InfoVis)*, *visual analytics (VA)*, *human-computer interaction (HCI)* and related fields to examine which and how immersive technologies can be used to improve data analysis and communication [56, 102, 343]. Thus, it extends the scope of visual analytics [366], for example, by employing technologies along the virtuality continuum. As defined by Milgram and Kishino [248], the continuum extends from the real environment via *augmented reality (AR)* to *augmented virtuality (AV)* and to *virtual reality (VR)*. Following this definition, unlike visual analytics, IA examines the impact of the technology used to remove the barriers between the analyst and the data for the exploration, interpretation, and understanding of data-driven problems [219]. In addition, it is not limited to visual representations of data but can use various stimuli, including sound and haptics.

In SciVis-related areas, the use of virtual environments for the visualization of spatial data has been common for decades – already, the first CAVEs in the early 90s were often dedicated to SciVis [42, 76]. IA employs many methods of SciVis and is used in various application areas such as archaeology [211, 351], geosciences [159], and the life sciences [79, 353]. Also in industry, CAVEs and other AR/VR environments are steadily gaining in popularity in various sectors, such as in healthcare, aviation, or automotive industry [384]. However, many challenges remain for IA to be effectively deployed across a range of application areas [108]. Especially applications with abstract data representations in 3D are considered problematic and often criticized in the visualization community [259]. Thus, further fundamental research is required to investigate the potential of immersive visualizations for data exploration and analysis [202].

Can novel technologies and methodologies help address previous criticism of abstract data visualization in 3D? How can visualizations such as network layouts, scatterplots, and parallel coordinates in immersive environments be designed to improve upon classic desktop setups? Are there convincing examples of such approaches?

To find answers to these and related questions, we evaluated the publications of 76 proceedings of eight conferences related to immersive analytics between 1990 and 2020 in this survey.

In this chapter, special emphasis is placed on visualization while excluding research that focuses on other, non-visual stimuli such as sonification [240, 424] or olfaction [40, 278]. Our focus here is on visualizations of abstract data, i.e., data without a natural physical or spatial representation, in immersive 3D environments, and we reviewed scientific literature that uses stereoscopic 3D for the inspection and analysis of visualizations. Hence, 3D visualizations that are exclusively inspected on 2D screens are excluded from our survey. Even with these restrictions, we could identify a significant increase in the number of publications over the last few years, indicating the need for and the interest in this research area.

Besides providing an overview of the field, our analysis revealed several interesting findings. While much literature on abstract visualizations and immersive environments exists separately, we found relatively few papers (48) that combine both fields by deploying immersive environments for abstract 3D visualizations. A number of interesting trends emerged, also with regard to the technology typically used. While for the 1990s and early 2000s, mainly CAVEs were found in our analysis, since 2017, HMD-related technologies dominate the publications, with a focus on VR, although AR is also an interesting area for IA research. The reason for this could be the still limited technical sophistication of AR (e.g., small field of view, limited interaction possibilities). Based on our analysis, we discuss the potential benefits of stereoscopic 3D visualizations, opportunities and challenges with regard to navigation and interaction in immersive environments, and the potential of immersive environments for collaborative analysis procedures in the context of abstract data analysis.

In the following section, we will first provide an overview of related surveys before describing our methodology and classification scheme in Section 2.3. The subdivision of this section provides the structure for all of the following parts. Thereupon we proceed to the actual core part in Section 2.4 in which we proceed through all analyzed dimensions. Each subsection is structured similarly. First, an overview of the distribution of all analyzed papers with regard to the respective dimension is provided. Subsequently, findings for each class of the dimension are presented while similarities and differences between approaches are highlighted. After revealing our results, high-level implications for immersive analytics are discussed in Section 2.5. In the final discussion (Section 2.6), we reflect on our findings and discuss various facets of immersive analytics with regard to abstract 3D visualization.

2.2 Related Work

This section provides an overview of related surveys that structure immersive analytics techniques according to various criteria and do not focus on individual application areas. The aim of this section is to provide the reader with a meta-analysis on immersive analytics and to illustrate the relevance of a systematic literature review on this topic.

Brooks was one of the first researchers to discuss positive and negative aspects of VR based on various applications [39]. In his early literature review, he came to the conclusion that despite the high cost, low resolution, and limited range of trackers, VR really works for specific domains such as flight simulators, automotive engineering, or astronaut training. However, some key features, like interacting with the virtual worlds or better modeling of the real world, remain challenging. One year later, van Dam et al. [374] highlighted VR applications for scientific visualizations. Examples that benefit from the integration of VR are in the area of archaeology for a better perception of ancient structures or in the medical field for a better understanding of the 3D geometry of blood vessels. Both works focus on the integration of virtual reality to replicate real-world scenarios or to display scientific visualizations. In this survey, we focus on *abstract data visualizations* for immersive analytics.

Laha and Bowman reviewed VR techniques for visualizing volume data [214]. In their literature review, they concluded that more controlled experiments are needed to explore the benefits of individual components of immersion. As a starting point, they came up with a task taxonomy that can later be used in user studies to generalize the results. Our survey is not limited to volumetric data sets but also includes relational and multidimensional data.

Reda et al. [304] specifically focused on summarizing research for hybrid reality environments like the CAVE2. Despite the advantage of a high-resolution screen in combination with optional stereoscopic depth, the authors emphasize the possibility of collaborative data analysis. We do not limit ourselves to hybrid reality environments but cover all technologies that make use of stereoscopic depth (e.g., CAVE, CAVE2, Volumetric Displays, HMDs).

Brath [34] collected evidence in the form of application examples that 3D visualizations offer advantages beyond 2D. Although the author does not focus on stereoscopic 3D, he mentions the benefits of an immersive interface. In contrast to his work, we explicitly restrict ourselves to immersive displays and not 2D screens.

The literature review by McIntire and Liggett [242] is closely related to the current one. The authors discussed the possible utility of stereoscopic 3D displays for information visualization. They focused on abstract data visualization and presented experiments in favor of and against stereoscopic 3D. We build on this work and additionally include application and evaluation papers from different domains.

García-Hernández et al. [128] focused on using virtual reality environments for visual data mining. Similar to our work, research on the representation of abstract visualizations like 3D scatterplots or 3D parallel coordinates in virtual reality environments is investigated. As opposed to their work, we do not limit ourselves to VR environments but also collect research in the field of augmented reality.

Sommer et al. [352] presented current research projects developed in collaboration between Monash University and the University of Konstanz. They concluded that stereoscopic 3D is advantageous in various application domains. While the authors presented seven research projects that make use of stereoscopic 3D, we do not restrict ourselves to specific projects but give a more comprehensive overview of research in the field of immersive analytics.

Just recently, Fonnet and Prié [120] surveyed 177 publications in the domain of immersive analytics and provide an excellent overview of different rendering technologies, data, sensory mappings, and interaction means which have been used to build IA systems. The biggest difference to the current work is that we restrict ourselves to abstract data visualizations and analyze corresponding papers in more detail by setting the focus on visualization types, analysis tasks, and the discussion of the applicability of abstract data visualization in immersive environments.

In summary, we are not aware of any previous work in which immersive analytics of abstract data for information visualization was systematically reviewed.

2.3 Methodology

In our systematic literature review, we focused on the selection of papers according to two different characteristics: *immersive environments* and *abstract data visualizations*. In order to be included in the survey, papers must use hardware that enables an immersive experience and employ techniques for visualizing abstract data. We consider all papers that meet both criteria but evaluate them in terms of their relation to immersive analytics. Although we made every effort to be accurate in our characterization of papers, there could, of course, be papers that are somehow related but still not included in our collection process (see Section 2.3.3). Given the current interest in immersive analytics research and presuming that this trend will continue, we can expect that further relevant papers will be published in the future. We, therefore, provide a customized online interface that allows extending our current survey by adding new publications on this topic and also provides interactive access to our collection: `https://iasurvey.dbvis.de`.

2.3.1 Immersive Environments: Sampling Characteristics

A prerequisite is that relevant papers are located in Milgram et al.'s virtuality continuum [249], which encompasses the entire range between the two extremes of the real and the virtual environment. The area between these two poles is called mixed reality, including augmented reality (where the virtual augments the real) and augmented virtuality (where the real augments the virtual). Therefore, the papers discussed in this chapter must be located in the domain of *immersive environments*, leading to a mixed reality experience. This means that abstract 3D visualizations must be presented in a mixed or virtual reality environment in which the hardware and user interact closely: The hardware monitors human behavior and reacts by stimulating human perception [218]. Thus, a paper that uses abstract 3D visualizations presented on a 3D projector is excluded by this restriction since no immersive environment in the classical sense is created, and there is no interaction between observer and systems. This means that head and body movements of the user have no influence on the perceived visualization, and the visualization is not fixed to a certain location in the visually perceived (real/virtual) environment of the user. Similarly, AR approaches are excluded if AR is purely created by handheld devices as head movements do not have any impact on the image perceived by the user - i.e., a virtual object depicted on the screen of a handheld device does not change perspective when the user looks on the display from different angles. This strict criterion also leads to the exclusion of approaches that create immersion with powerwalls.

In order to search for all contemplable papers, we created a keyword list with terms related to immersive environments, such as '3D', 'virtual reality', 'augmented reality', and 'immersion' (see the complete list in Appendix II). These terms were compiled from the experience of the authors as well as from the typical jargon of previously known prominent literature in this domain like the work of Slater and Wilbur [348]. All keywords were preprocessed using standard natural language processing algorithms such as stemming to increase the chance of a positive match. As a result, we collected a first keyword list with 38 keywords covering the concept of *immersive environments*.

2.3.2 Abstract 3D Visualizations: Sampling Characteristics

Abstract data can be defined as data that has no inherent spatial structure or physical representation [103]. Abstraction in visualization is achieved through the use of color and shapes that are not directly related to the object in question [285]. We investigate publications dealing with abstract data visualizations. Therefore, we introduce three inclusion criteria that were used to filter the large number of immersive analytics papers. In this way, we consider three categories:

(c1) Visualizations of abstract data (i.e., data that has no inherent mapping to space, including all visualizations of abstract data).

(c2) Visualizations of abstract data in a spatial context that are situated representations [409] (i.e., the abstract visualization element displays data in proximity to data references but does not spatially coincide with data references).

(c3) Embedded visualizations of abstract data in a spatial context that encode more than one attribute with visual variables (e.g., glyphs, space-time cubes).

At the same time, we exclude papers with visualizations from the following categories:

- Pure non-abstract data visualizations (e.g., a 3D visualization of a brain, an engine, or a map).
- Embedded visualizations [409] of abstract data in a spatial context that use only a single visual variable (i.e., visualizations in which the displayed data match data references, such as colored blood vessels in a 3D brain model, text labels associated with a 3D engine visualization, a 3D map with colored dots representing specific locations).

Literature that introduces abstract data visualizations must specify either the data type as abstract or the visualization technique used to display the respective data. Therefore, we focused on both aspects when we created a second keyword list for filtering contemplable papers which contained keywords such as 'scatterplot', 'abstract data', 'high-dimensional', and 'PCP' (see the full list in Appendix III). Abstract data has no inherent spatial structure; examples are hierarchies, networks, or multidimensional data points. Prominent visualization techniques include scatterplots, parallel coordinates, or pixel-based visualizations. To draw up a comprehensive list of adequate visualizations, we reviewed the literature implemented in lectures on information visualization. Popular examples are the books "Interactive Data Visualization: Foundations, Techniques, and Applications" [396] or "Visualization Analysis & Design" [259]. Again, all keywords were preprocessed to avoid the influence of affixes on matching terms. As a result, we collected a second keyword list containing 79 keywords covering the concept of *abstract data visualizations*.

2.3.3 Paper Sampling

We parsed the proceedings of the most important conferences in the field (i.e., BDVA, CHI, ERVR, EuroVis, IEEE VIS, IEEE VR, SD&A, UIST - see Appendix I) and applied a full-text keyword search. The keywords were chosen to be descriptive for *immersive environments* and *abstract data visualizations* (see section 2.3.1 and section 2.3.2 - list of keywords in Appendix II and III). Whenever a combination of keywords from both categories was found in the text, the respective paper was selected as a potential candidate. After this automatic parsing and matching process, we created a candidate corpus of 256 papers. The initial selection of papers was intentionally given weak constraints in order to not exclude relevant papers and to create an extensive pool of candidates. Due to these weak constraints, the pool contained many false positives that did not meet our predefined criteria. Therefore, the set of 256 papers was reviewed and filtered in a manual screening process based on our strict definition of *immersive environments* (section 2.3.1) and *abstract data visualizations* (section 2.3.2). We also identified surveys and state-of-the-art reports and excluded them from further analysis. Such reports

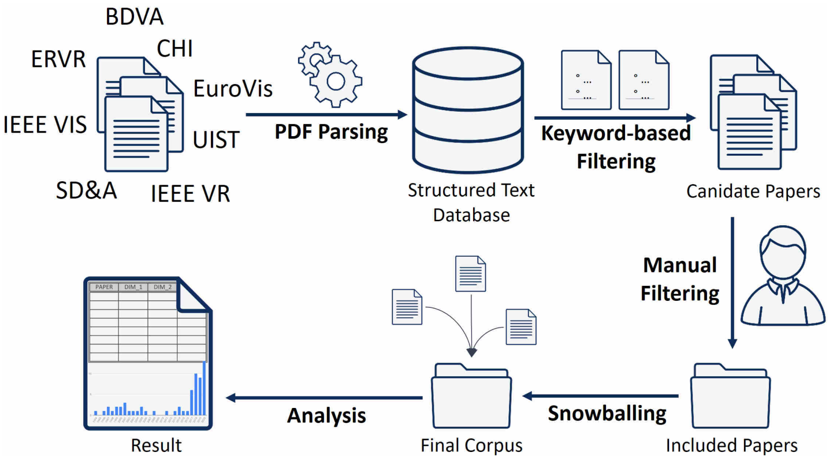

Figure 2.1: The sampling process of our survey is based on parsing and keyword filtering of a large set of papers in PDF format. After a subsequent manual filtering step, the initial set of included papers is generated which is then used as a starting point for the final expansion by manually parsing the reference lists of included papers (snowballing).

would distort a detailed analysis because they do not focus on a single new approach but describe several techniques in a single paper. The overall sampling process is depicted in Figure 2.1.

To give an example: the title of the paper "Objective and subjective assessment of stereoscopically separated labels in augmented reality" [283] sounds promising in the context of our review. However, after reviewing the paper, it turned out that a) it does not apply immersive analytics principles and b) the labels mentioned in the title identify airplanes rendered as simple 3D objects – thus "representing embedded visualization of abstract data in a spatial context that only uses a single visual variable". Therefore, this paper was excluded. Another example is the work of Greffard et al. [139]. Although the abstract data visualization criteria are very well suited for this work as it evaluates graph visualization, it does not fit the second criterion, which concerns the immersive environment: Although a stereoscopic screen is used for the visualization, the paper uses a static visualization, and no head-tracking is involved - hence, the degree of immersion is relatively low. Our manual screening procedure resulted in a set of 28 papers (our basic set) and served as a starting point for further paper acquisition.

To broaden the scope of our initial semi-automatic sampling strategy, we used a snowball sampling technique [414]. More precisely, we recursively scanned the references from all papers in our basic set and checked them for relevance. Using this approach, we collected another 20 papers in two iterations, so that a total 48 papers were subjected to our detailed review process. With this approach, we were able to cover a wide range of journal papers in addition to the originally parsed set of conference proceedings. Some papers found during the recursive parsing procedure were not detected during the semi-automatic sampling because they were not included in the paper pool (different venues, excluded years), because parsing errors led to mismatching keywords, or because no or only one keyword was used in the paper. We carefully tried to optimize the PDF parsing process and did our best to identify papers with parsing errors in order to scan them manually for relevance. However, with a set of in total over 19,400 papers, it is impossible to guarantee that not a single paper with parsing errors was overlooked.

2.3.4 Analyzed Characteristics

In the following, we will introduce the classification used to group and organize the set of inspected papers. We classified the papers according to six characteristics: Paper type, technology, environment type, data type, visualization technique, and analysis task. The classifications of paper type and data type were adopted from Munzner [258], whereas the classifications of the other characteristics were derived in a bottom-up approach from the inspected literature. All papers were assigned to one or more classes per characteristic. Only for paper type, each paper was assigned to exactly one class.

2.3.4.1 Visualization Technique

Starting from categories found in visual analytics-related literature like in Munzner [259], we identified seven types of visualization techniques used in the investigated set of papers. Additionally, we grouped techniques appearing only once into the supplementary category 'others'.

Node-Link Graph
Networks and tree visualizations.

Scatterplots
Multiple data entries, represented as points in 2D/3D coordinate systems.

Parallel Coordinate Plots (PCPs)
Multiple data entries, represented as lines between arranged axes.

Glyphs, Icons & Symbols
Visual data metaphors that often encode more than one dimension.

Geographic
Real-world geometry representations.

Volume
3D object visualizations.

Flow
Scalar-, vector-, and tensor field visualizations.

Others
Rare techniques that do not fit into any of the above categories, such as height map visualizations or Kohonen map representations.

2.3.4.2 Analysis Task

Based on Andrienko and Andrienko's task taxonomy [4], we classified the papers into higher-level (synoptic) tasks and elementary tasks. For a more detailed analysis, we further distinguished seven categories of analysis task if the respective task was mentioned as a valid analysis task for the proposed technique. Papers that are not assignable to any of the seven classes or do not explicitly state an analysis task are categorized as 'other' or 'not specified', respectively.

Synoptic tasks:

Clustering / Classification
Structuring and grouping data points.

Anomaly Detection
Finding anomalies, such as outliers, in datasets.

Pattern Analysis
Finding trends, repetitions, and visual patterns in datasets.

Visual Search
Often used in exploratory studies to visually identify and track an object or data point.

Overview
Providing the big picture of a dataset, often in combination with details on demand.

Elementary tasks:

Comparative Analysis
Comparing different datasets, such as trajectories or molecular structures.

Data Enrichment
Adding information to the data, such as interactive label placement.

2.3.4.3 Paper Type

All papers were assigned to exactly one of the following paper types (adopted from Munzner [258]).

T *Technique*
Papers presenting novel algorithms and techniques.

E *Evaluation*
Papers with the focus on the assessment of an application, approach or technique.

S *System*
Papers describing the architecture of a framework.

M *Model*
Papers providing a theoretical view of things.

D *Design Study / Application*
Papers presenting the application of existing techniques to solve a certain problem in a certain domain.

2.3.4.4 Technology

Immersive technologies were categorized into three groups, which cover all technologies deployed in the set of papers under consideration.

Monitor / Projector
3D monitors or projectors used to create semi-immersive environments.

CAVE
Video-wall environments of different shapes.

HMD
Head-mounted displays worn by a user to enter AR or VR environments.

2.3.4.5 Environment Type

Since a prerequisite for the inclusion of papers was the embedding of the approach on the virtuality continuum, and we are not aware of any paper making use of augmented virtuality under given conditions, the resulting corpus contains only AR and VR papers.

AR
Augmented Reality - virtual objects are embedded into the real environment.

VR
Virtual Reality - a purely virtual environment is perceived.

2.3.4.6 Data Type

The datasets used were categorized into four classes (adopted from Munzner [258]).

Table
Items in a table refer to individual data points, whereas attributes or dimensions refer to the data dimensions of the data points. The combination of an item and an attribute is reflected in a single cell containing a value.

Field
In a continuous domain, fields represent attribute values associated with cells. The resolution may change depending on whether the density measures are closer together, which leads to a higher resolution, or whether they are further apart in a coarser grid.

Geometry
Geometric items can be points, lines, curves, 2D surfaces, or 3D shapes with an explicit spatial position. Geometric datasets can come with additional attributes, making their visualization a challenging task, or without.

Network
Data points that have a relationship to each other can be specified in the abstract concept of a network with nodes and links. These nodes and links can be associated with attributes specified in tables. In this survey, we do not distinguish between networks with cycles and hierarchical structures.

2.4 Literature Review

In the following subsections, we present our results for each of the six dimensions considered in the same order as introduced in the previous section. In each subsection (dimension), we will first provide an overview on the dimension itself and its development over time to then have a close look at all its classes. The classes are categorical and, therefore, presented without a specific order. Papers within each class are discussed in semantic, then chronological order. For an overview of all analyzed papers, we provide a link to an online browser that offers advanced search, filtering, and comparison options: `https://iasurvey.dbvis.de`. The online platform not only allows the overview to be expanded to include future work but also missing papers or even missing dimensions to be added. The overall distribution of papers considered in this survey is shown in Figure 2.2. While research interest was high in the late 1990s and early 2000s, we observed a decline in research papers in the late 2000s and early 2010s. This could be explained by increased research efforts regarding the application of immersive environments for non-abstract immersive visualizations and a generally declining research interest in abstract data. In recent years we notice a strong research trend towards abstract 3D visualizations in immersive environments. This could be due to the steady progress of technology and improved availability.

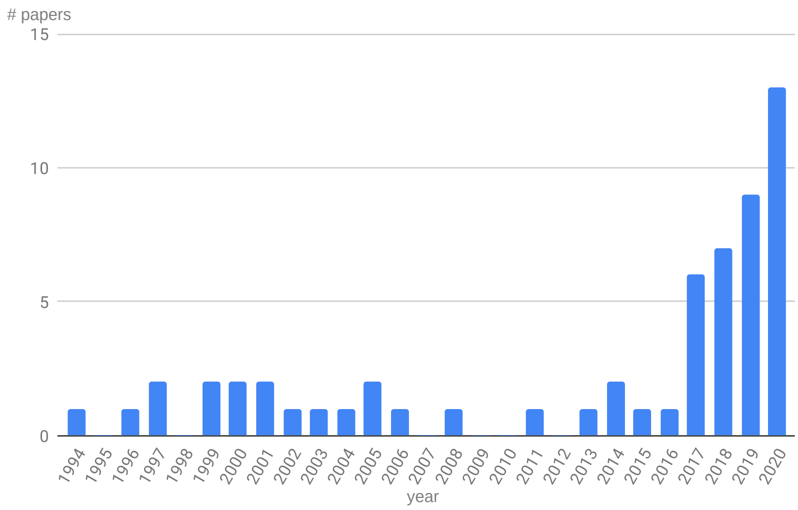

Figure 2.2: Literature review: Distribution of analyzed papers over time.

2.4.1 Visualization Techniques

While a variety of visualization techniques has been developed to meet different requirements regarding data types, visualization aims, and tasks, these techniques also have their own requirements, affordances, and restrictions regarding the environment in which they are used and the associated interaction operations. As a result, the potential design space for their integration, as well as the effectiveness and efficiency of the techniques, can vary depending on the immersive environments and devices. Related aspects such as field of view and field of regard, resolution, screen size, and computational power may affect the suitability accordingly. A considerable number of papers employ multiple visualization techniques, often with different levels of support and description, so multiple entries are possible (see Table 2.2).

Table 2.1: Summarized results of the literature review.

	Paper Type	Technology	Environment	Data Type	Visualization Technique	Analysis Task
[398]	E		VR			
[399]	E		VR			
[364]	D		VR			
[377]	D		VR			
[6]	E		VR			
[266]	E		VR			
[181]	D		VR			
[328]	S		VR			N/S
[349]	T		AR			
[261]	S		VR			N/S
[360]	D		VR			
[18]	E		AR			
[301]	E		VR			
[16]	E		VR			
[115]	D		VR			O
[9]	D		VR			O
[260]	S		VR			
[405]	S		VR			
[111]	E		VR			
[96]	E		VR			
[152]	D		VR			
[354]	T		AR			N/S

Table 2.1: Summarized results of the literature review (cont.).

	Paper Type	Technology	Environment	Data Type	Visualization Technique	Analysis Task
[213]	E		VR			
[19]	T		VR			
[75]	E		VR			
[116]	E		VR			
[323]	D		AR			
[74]	S		VR			N/S
[26]	M		VR			N/S
[48]	T		AR			
[98]	E		VR			
[121]	E		VR			O
[236]	D		VR			
[15]	D		VR			
[255]	T		VR			
[168]	D		VR			
[341]	S		VR			N/S
[166]	S		VR			
[422]	E		VR			
[389]	E		VR			
[199]	T		VR			
[143]	S		AR VR			O
[209]	E		VR			O N/S

Table 2.1: Summarized results of the literature review (cont.).

	Paper Type	Technology	Environment	Data Type	Visualization Technique	Analysis Task
[73]	S	Hmd	VR	Table	Scatterplot, PCP	N/S
[208]	E	Hmd	VR	Table	Scatterplot	Cluster
[311]	D	Monitor, Hmd	AR	Table	Scatterplot, PCP, Other	Enrichment, N/S
[198]	T	Hmd	VR	Field	Other	Comparison
[332]	E	Hmd	AR	Table	Other	Search, Overview
[71]	T	Hmd	AR	Table	Scatterplot, Other	N/S
[149]	T	Hmd	VR	Table	Other	Overview, N/S
[428]	T	Hmd	VR	Table	NodeLink, Glyph	Search, Overview
[227]	E	Hmd	VR	Table	Other	Comparison
[403]	E	Hmd	AR, VR	Table	Scatterplot	Anomaly, Pattern, Search
[158]	D	Hmd	VR	Table	Geographic	Search, Overview
[265]	S	Hmd	AR	Table	Scatterplot, Other	N/S
[47]	S	Hmd	AR, VR	Table	Geographic, Scatterplot, PCP, Other	N/S
[221]	E	Hmd	VR	Table	Scatterplot, Other	Search, Overview
[421]	E	Hmd	VR	Table	Scatterplot	Overview

Paper Type	Technology	Environment	Data Type	Visualization Technique	Analysis Task
T Technique, E Evaluation, S System, M Model, D Design Study	Monitor, Cave, Hmd, Tablet	AR Augmented Reality, VR Virtual Reality	Table, Field, Geometry, Network	NodeLink, Scatterplot, PCP, Glyph, Geographic, Volume, Flow, Other	Cluster, Anomaly, Pattern, Search, Enrichment, Overview, Comparison, O Other, N/S Not Specified

Table 2.2: Literature review. Visualization techniques.

Visualization Technique	References
Node-Link Graph	[18, 26, 75, 98, 115, 149, 209, 213, 236, 349, 354, 360, 398, 399, 428]
Scatterplots	[6, 9, 26, 47, 71, 73, 74, 96, 111, 116, 121, 208, 221, 260, 261, 265, 266, 301, 311, 341, 364, 377, 403, 421]
Parallel Coordinates	[26, 47, 48, 73, 74, 311]
Glyphs, Icons & Symbols	[19, 377, 428]
Geographic	[15, 47, 152, 158, 323, 341, 389, 422]
Volume	[143, 168, 236, 255, 341, 354, 377]
Flow	[16, 158, 166, 181, 255]
Other	[47, 71, 149, 198, 199, 221, 227, 265, 311, 332, 341, 405]

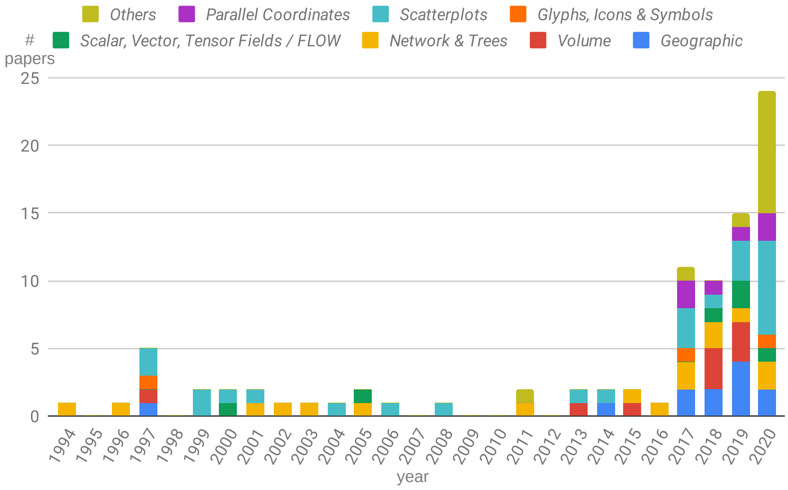

Figure 2.3: Literature review. Distribution of visualization techniques over time.

Our analysis of these visualization techniques shows an unbalanced distribution, with a clear focus on two types of techniques under investigation: node-link graph and scatterplot, which comprise 39 out of 48papers. As can be seen in Figure 2.3, there has been a sustained interest in these techniques for many years, while the investigation of geographic visualizations has recently experienced a boom. While text visualization is a natural component of many visualization approaches, e.g., for labeling, and there are approaches to analyze text corpora, we have not found any paper that focuses on a text visualization technique for immersive analytics approaches. This may be related to the fact that for many of the immersive technologies, such as VR and AR HMDs, current devices have some shortcomings in terms of dynamic text rendering due to relatively low resolution and the additional impact of the distance and perspective of 3D objects, which makes it difficult to create high-quality text visualizations. Various papers contained volume and geographic visualizations but were excluded due to insufficient levels of abstraction. For example, papers were excluded that only visualize 3D models of blood vessels or present 3D geo-maps without encoding additional information. A few papers did not fit into a clear-cut classification or would form a single-element class of their own, such as a SOM-based visualization approach for multidimensional data [405] or heightmap visualizations [199], and were therefore subsumed under 'Other'. In the following, we take a close look at each class of visualization technique.

2.4.1.1 Node-Link Graphs

Since they lack predefined axes, dimensions, and directions, network visualizations offer much freedom in creating visual representations. However, this also means that there is usually a less unified user experience, and many design decisions can distort the validity and effectiveness of solutions. In addition, there are a variety of characteristics in networks created from application data, including scale, but also structural features. Techniques might be suitable only for very specific subclasses of networks, and the impact of the immersive environment might further strongly influence usability. Therefore, the practical evaluation of techniques in user studies is of utmost importance. On the other hand, the potential of visual network analysis in immersive environments has already been demonstrated in a number of studies that focus on aspects such as improved collaboration and interaction, better perception of network features through stereoscopic views, and visual scalability (see following examples). With the ever-increasing size and complexity of datasets, the question of how to support the human mental map for navigation in network visualizations is also gaining importance in current research.

Perception & Human Factors — In their seminal work on stereoscopic 3D perception of networks, Ware and Franck [399] evaluated the influence of stereoscopic 3D visualization and motion cues compared to 2D visualization under various conditions, including head tracking, in a setup with shutter glasses and fish tank VR. Expanding on their earlier report [398], they found clear improvements when the head-coupled stereo condition was applied, but also argued that the type of motion applied, for example, automatic rotation, should depend on the application and the required interactions. In view of the special setup and the relatively small size of the random networks used, the results must be checked for more general evidence. Belcher et al. [18] investigated the use of AR for the analysis of complex networks and provided a user study based on artificially created networks. They reproduced the classical experiments by Ware and Franck [399] and compared AR with 2D and 3D screen settings using simple node-link visualizations. They conclude that the limitations of the AR technology, e.g., regarding color and contrast, might still hamper its effectiveness in task performance. Given the improvement in technology in recent years, this hypothesis could be reevaluated with current technology. While some investigate human factors and perception on network visualizations in oder to improve them or associated tasks, others simply make use of graph visualizations as a means for general assessments on perception or human factors. For instance, Krekhov et al. [209] evaluated the deadeye highlighting technique, besides on others, on network visualizations.

Navigating Graphs — In a comparison of two classes of environments, CAVE and VR HMD, Cordeil et al. [75] compared the task performance in graph analysis tasks, triangle counting, and shortest path finding, for a collaborative setting with two participants in a team. While participants were faster in the HMD setting and movement differed between team members in the CAVE setting, no other significant differences were found in the collaborative task solving, which shows the potential of the VR HMD technology for such setups. Using VR HMDs with hand-held controllers, Drogemuller et al. [98] evaluated the task effectiveness of four navigation techniques for graph analysis, with one- and two-handed flying was perceived by participants as faster and more preferred than teleportation in search tasks (see Figure 2.4). Slay et al. [349] used augmented reality with fiducial marker-based interaction for the visual analysis of trees and graphs to demonstrate new object manipulation techniques.

Layouting Graphs — Kwon et al. [213] investigated the use of a spherical layout to improve network perception and interaction in VR and performed a comparative analysis of 2D and 3D graph layouts. They found that in their setup with networks of up to 297 nodes and 2359 edges, participants

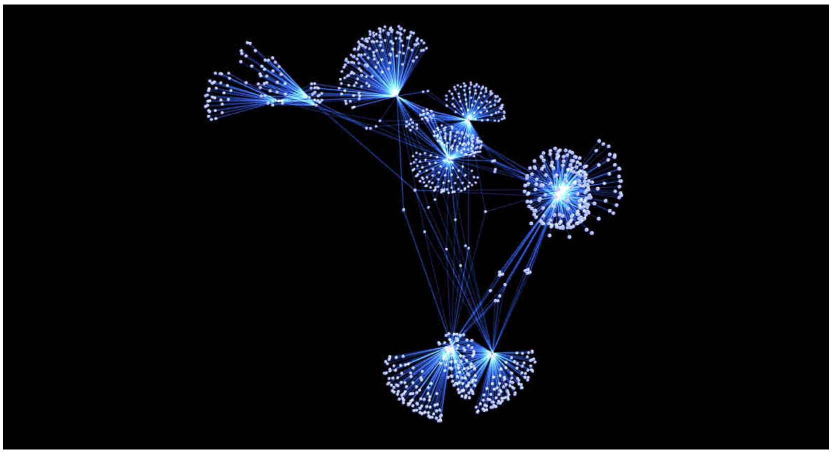

Figure 2.4: 3D network visualization of abstract data in an immersive VR environment [98]. Image courtesy of Adam Drogemueller.

solved tasks with the spherical layout and the corresponding interaction technique significantly faster and with a significant increase in correct answers for larger graphs.

Analytic Provenance & Processes as Graphs — Besides such typical approaches that are concerned with the visualization of network or hierarchical data, other visualizations exist that fall into the same category of 'Node-Link Graphs'. For instance, Hayatpur et al. [149] present a visualization approach for analytics provenance graphs, which are generated throughout analysis procedures. Even though this is not a typical graph visualization problem, it still is, in principle, a node-link graph visualization of abstract data. Similarly, the approach by Zenner et al. [428], in which abstract process models are transformed into interactive 3D environments, resemble node-link visualizations.

Domains — Since biology is one of the most important application areas for network analysis and visualization, a major research focus is on techniques that take into account the specifics of the corresponding datasets, tasks, and notations. In particular, the flood of data resulting from high-throughput 'omics' technologies, e.g., for proteomics and genomics analysis, and the resulting need for methods that can cope with the scale and complexity are a driving factor for current research. Ferey et al. [115] described the Genome3DExplorer for the investigation of genome data. Networks are used to model binary relations between genomic entities, such as yeast gene coexpression, and provide an interactive 3D visualization making use of a force-directed layout approach. Maes et al. [236] presented MinOmics, an analysis pipeline for multi-omics data, and discussed several scenarios for using interactive, immersive environments such as stereoscopic display walls and VR HMDs while describing implementation work in progress in this direction. Stolk et al. [360] presented an approach to mine genomics data in which relationships between entities are represented as node-link representations in 3D VR without focusing on a specific environment or device. Due to the density of the resulting networks, they resort to edge filtering based on similarity values. As abstract data is often associated with spatial data, representation, and navigation in such cases must take both data types into account. Sommer et al. [354] investigated a combined 2D and 3D approach for navigation and demonstrated it with an application on cytological network exploration that links network visualizations to 3D cell model rendering.

2.4.1.2 Scatterplots

Scatterplots and scatterplot matrices are well-known visualization techniques for the analysis of high-dimensional data in 2D. Starting in the late 90s, researchers investigate their performance in VR environments. With that, the visualization technique gained more and more in popularity - also as a secondary tool for evaluating human factors or user experience with largely independent interaction techniques or hardware. In total, we analyzed 24 papers that make use of scatterplot visualizations.

Immersion, Visual Perception & User Experience — Although many use cases were introduced, experimental evidence demonstrating the benefits of 3D scatterplots in VR is rare. Arns et al. [6], and Nelson et al. [266] investigated participants during a cluster identification task in a VRE and on a desktop monitor. Results showed that participants in the VRE performed almost twice as well but needed a little more time to become familiar with the interaction possibilities. The authors attribute this to the display of "true" three-dimensionality and the improved perception of structures in the VRE. However, the question of whether the benefits are due to the additional third dimension or to the fact that analysts are more immersed in the analysis is still pending. To study the effect of physical engagement in an immersive environment, Raja et al. [301] conducted an experiment using a CAVE with one and four walls. Participants had to analyze data in a 3D scatterplot. Additionally, the authors enabled head tracking for one group of participants to increase the level of immersion. Results suggest that participants are more efficient in a highly physically immersed environment such as a CAVE with four walls and head tracking. The VRMiner tool [9] makes use of scatterplots for interactive mining of multimedia data in VR and maps visual variables to display different types of data in an abstracted form. The authors argue for the benefit of the deployed VR environment due to the easier perception of the presented data and direct interaction capabilities. However, the authors did not perform a direct comparison comparing the performance or usability of their VR solution with a conventional screen-based setup. Filho et al. [116] compared screen-based setups with HMDs for visual exploration tasks on multidimensional data represented as scatterplots. Their study results showed beneficial effects of immersion in terms of distance perception and outlier identification tasks. In addition, they demonstrated higher accuracy and engagement scores when participants were in the immersive environment. Similarly, Whitlock et al. [403] used scatterplots to investigate a possible difference in the perception of visual variables between different media (AR, VR, Screen).

Orienation & Navigation — Etemadpour et al. [111] compared different segregation and precision tasks performed in VREs and non-immersive 2D environments on the screen. The three-dimensional data were displayed as 3D scatterplots. Correctness, timing, and confidence were higher for several tasks when participants were in the immersive environment. Among other things, they were able to show that participants were able to approximate distances better when they were in the virtual environment. However, they reported a loss of orientation when participants were in the immersed environment. This loss of orientation was also identified by Kraus et al. [208] in their controlled user study. They propose not to surround users with data points but to provide a restricted area in the VRE where data is displayed. Such an overview prevented the aforementioned loss of orientation during a cluster identification task. In comparison to more abstract representations on 2D screens, the VRE increases memory and orientation capabilities by providing more natural navigation in the data space.

Interaction — First prototypes of 3D scatterplots were introduced by Symanzik et al. [364] and Teylingen et al. [377]. Both approaches made use of similar interaction techniques like rotating and moving the visualization space and interactive menus to select or filter data. However, there was no evidence whether such a third dimension really improves the analysis of scatterplots. In the following years, the design and interaction space of 3D scatterplots was improved with scatterplot matrices [261], more sophisticated interaction techniques like details-on-demand [260], or possibilities to easily change visual parameters [96] (Figure 2.5). Furthermore, Donalek et al. [96] implemented a multi-user setting in their prototype to support the cooperative analysis of data. This was achieved with a broadcasting function where a user shares his view on the data with his colleagues.

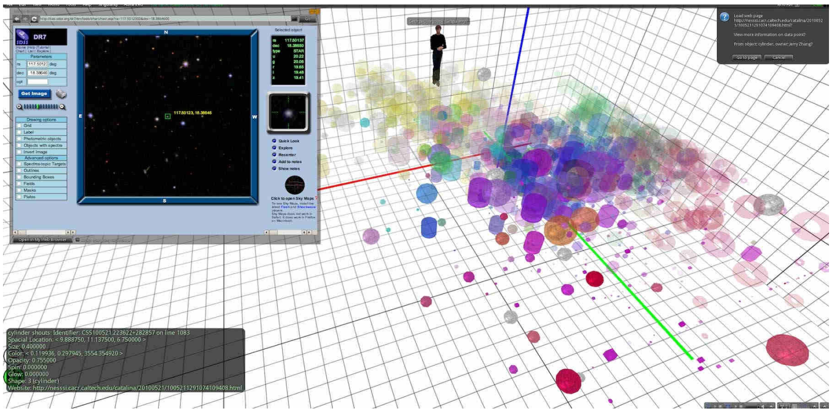

Figure 2.5: This image shows a student represented by an avatar performing experiments in an 8-dimensional data visualization: Data parameter values are mapped into the displayed 3D space by data point shapes, sizes, colors, and transparencies, representing an 8-dimensional data visualization [96]. Image © 2014 Institute of Electrical and Electronics Engineers.

***Design Space* —** To investigate the design space of scatterplot axes, Fonnet et al. [121] conducted an experiment comparing plane axes in a 3D scatterplot with a semi-transparent 3D grid and plane grid lines with additional visual cues to facilitate orientation. Although participants were slower in the tasks with the 3D grid, most of them preferred this variant. Interestingly, there was no advantage in including additional visual cues as reference markers.

***Toolkits & Prototyping* —** To further facilitate the generation of 3D scatterplots, Cordeil et al. [74] introduced ImAxes, a visualization authoring toolkit for visualizing multidimensional data in VREs. Dimensions from the data can be interactively positioned in the environment as axes. Thus, analysts could easily add, rotate, or combine axes to create scatterplots, scatterplot matrices, or other kinds of visualization techniques. Likewise, Sicat et al. [341] implemented an approach to create immersive visualizations like 3D scatterplots. Their toolkit comprises a simple grammar and provides reusable templates for easy creation of unique 3D visualizations.

2.4.1.3 Parallel Coordinates

Parallel coordinate plots (PCPs) are an established technique that strives to overcome the limitations of scatterplots for high-dimensional data and to support traceability across all dimensions for a data point. While 3D might help to alleviate the problem of finding the correct order of dimensions, as only correlations between adjacent dimension are clearly perceptible, there is only very limited work on PCPs for IA, mainly describing techniques that are available in software implementations without a deeper analysis of the benefits and the potential in the IA design space.

***Dynamically Layouting PCPs* —** The previously mentioned visualization authoring toolkit ImAxes [74] allows the arbitrary arrangement of axes and with that the creation of PCPs and 3D

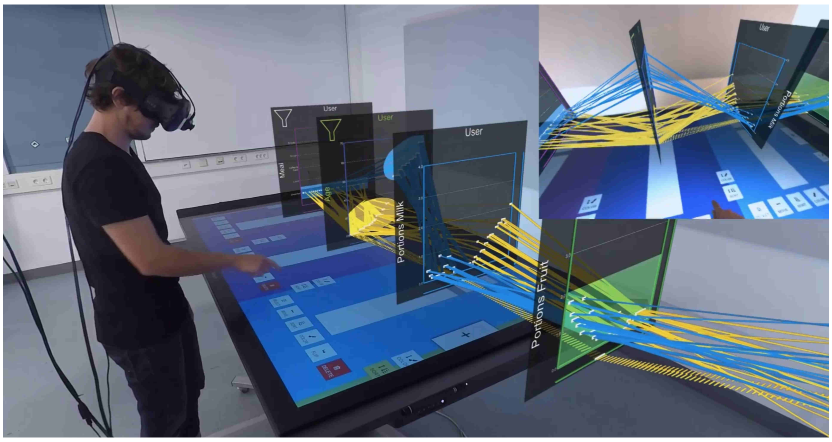

Figure 2.6: A 3D parallel coordinate plot (PCP) visualization of abstract data observed in a collaborative AR environment [48]. Image courtesy of Simon Butscher.

variants thereof - such as circular arranged ones or 3D PCPs where 2D scatterplots are connected with lines. Similarly, GeoVisor, a system presented by Billow et al. [26], allows users to visualize data in various ways - including as PCPs. However, unlike ImAxes, GeoVisor is not optimized for interactively authoring new visualizations but provides a sample framework for evaluation. Along with the proposed system, they demonstrated its evaluation based on heuristics.

PCPs in Hybrid XR Applications — Butcher et al. [48] used AR to visualize 3D PCPs on top of a touch-table and investigated the usability of the combination of AR and a 2D touch-table display (see Figure 2.6). They concluded that their approach facilitates immersion in the data, fluid analysis processes, and collaboration. They argue that the combination of AR with touch input could improve usability due to familiar, precise, and physically undemanding touch interactions compared to gesture-based interaction capabilities typically provided in AR.

2.4.1.4 Glyphs, Icons & Symbols

In proportion to the relative number of general research papers published for techniques based on glyphs, icons, and symbols, there is also a relatively small number of papers discussing these techniques in the context of immersive analytics. A particular challenge here is how these simplified representations can be lifted into the immersive design space to benefit from the extended possibilities without losing the advantage of simplicity, e.g., when exploiting stereoscopic 3D vision.

Glyphs in VR — An early example how this could be accomplished is presented by Teylingen et al. [377] presented the “Virtual Data Visualizer”, an immersive VR environment for visualizing data points as customizable glyphs. The system deploys traditional menus within the VR environment in combination with direct icon manipulations in 3D to allow users to create and customize glyph visualizations by manipulating mappings between variables and glyph elements.

Virtual Environments as Glyphs — More recently, Bellgradt et al. [19] presented ‘Gistualizer’, a tool for visualizing single, multidimensional data points as “immersive glyphs”. A landscape is automatically generated from the properties of the data point, with different dimensions defining the appearance of the landscape. For instance, one attribute defines the number of houses, another the height of mountains. Figure 2.7 depicts immersive glyphs created with Gistualizer, where one dimension varies between the five data points, affecting the depicted time of the year in the respective glyphs. The visualizations of abstract process models in the form of interactive 3D environments presented by Zenner et al. [428] can also be seen as large glyphs. Similar to the approach in "Gistualizer", the environment is automatically generated based on an abstract data foundation and can be explored by the user.

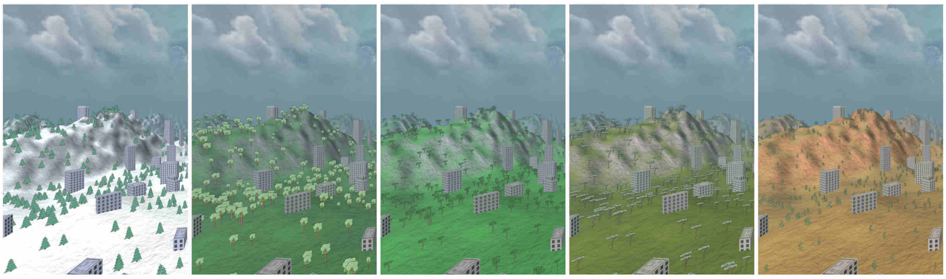

Figure 2.7: Different immersive landscape glyphs created with Gistualizer [19]. The appearance of the virtual environment itself represents one single data point. Hence, the created visualization resembles a glyph representation of abstract data. Image courtesy of Martin Bellgardt.

2.4.1.5 Geographic

Geographic visualizations are among the earliest data representations due to their use in exploration, navigation, urban planning, and agriculture. They involve geospatial referencing information and therefore often use map-based representations that are complemented by other abstract data representations such as availability or consumption of resources [95]. In particular, these visualizations can take advantage of immersive environments in terms of available space and navigation capabilities compared to traditional desktop settings. Since the benefit of specific data representation concepts, especially including stereoscopic 3D representations and interaction, can vary greatly between classic desktop settings and immersive environments, many concepts are revisited to explore the new possibilities.

***Maps* —** To this end, Yang et al. [422] visualized flows on maps, explored the design space and compared several VR representations of flow and the reference space. The design of geographic visualizations leaves certain freedom in the choice of view (egocentric vs. exocentric) and map representation (curvature and projection method). These can have a substantial impact in immersive environments, e. g., in terms of perceived distortion, and therefore both theoretical investigations and user studies are required for an assessment.

***Climate Visualization* —** A promising application area for which the use of stereoscopic 3D visualizations in immersive environments can be further explored is the investigation and presentation of climate data. Important questions include how data dimensions such as temperature, heat flux, precipitation, or humidity can be presented in a 3D visualization of the environmental context, how they can be combined without visual occlusion and overload, and how dynamics can be integrated. Helbig et al.[152] presented a workflow for the integration of heterogeneous data from simulation model variables and observed data with topographic features and other data about the environment in which they are situated. The workflow is designed for a VR environment and demonstrated in a projection-based stereoscopic virtual environment. Baltabayev et al. [15] addressed a similar problem, the visualization of collected environmental data from sensors deployed in the environment, and demonstrated a concept based on the reconstruction of the real environment in VR.

***Space-Time-Cubes* —** The space-time-cube (STC) is one of the standard visualization approaches for spatio-temporal geographic data in which lines are drawn within a cube and their location encodes a geolocations (x+z) and time (y/height). The technique was also ported to immersive environments. For instance, Saenz et al. [323] investigated the utility of immersive 3D visualizations for geographic data. They deployed AR holograms of space-time cubes and proposed a study design for future investigations. Similarly, Wagner et al. [389] presented a user study for an immersive STC implementation using gestures and tangible controls for interaction as well as a desk-based metaphor instead of flying or physical walking (see Figure 2.8). According to them, their study results indicate clear qualitative benefits for the exploration of trajectories with immersive STCs.

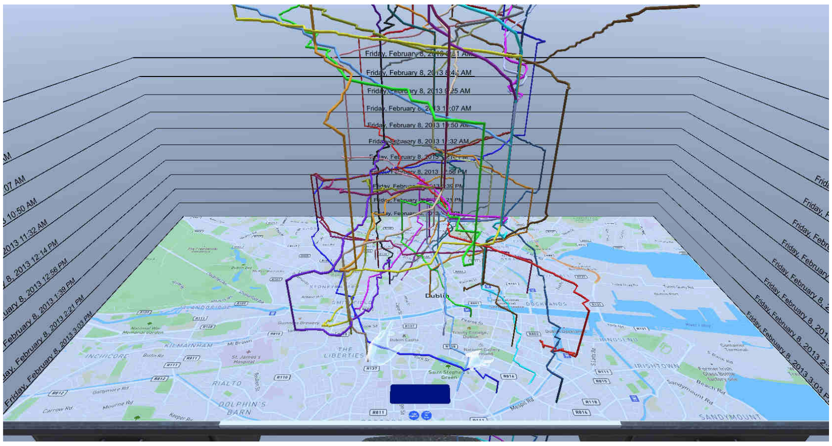

Figure 2.8: A space-time cube is visualized in a virtual reality environment by Wagner et al. [389]. The abstract variable time is mapped to the height of trajectories embedded in the geographic visualization. Image courtesy of Jorge Wagner.

2.4.1.6 Volume

The visualization of 3D volume data has attracted much attention in scientific visualization research. We identified seven papers that investigate the use of volume visualization with abstract visualization elements within immersive analytics approaches.

Application Domains — Sommer et al. [354] used a semi-immersive display to explore a virtual cell environment and a car model interactively. An embedded biological network structure was integrated into the cell and connected to its different components. For navigation through the car model, a labeled 2D map could be used to move the 3D view to the corresponding car components interactively. Similarly, Maes et al. [236] presented a tool that combines the visualization of molecular structures via VR headsets in UnityMol with omics-network visualization and analysis. For example, a Redox PTMs 3D network in C. reinhardtii was visualized and could be explored side by side with the visualization of the corresponding protein complexes. Still in the biology domain, Gunther et al. [143] introduced a Java framework for VR/AR bio- visualization that can process mesh and large volumetric data with multiple views, points in time, and color channels using OpenGL and Vulkan rendering APIs. This work presented a simulation of 10,000 agents that together form a sphere as well as an out-of-core 500 GiB multi-timepoint embryo dataset. DXR [341] is a Unity-based toolkit for creating immersive environments using concise declarative visualization grammar using the in-situ GUI. DXR's visualization pipeline supports templates and customizable graphical marks, which can be used to specify unique and engaging visualizations. DXR infers missing parameters to reasonable defaults and uses the inferred specifications to construct a 3D visualization that can be placed in a VR scene. A main focus of DXR is the visualization of abstract data by using 3D flow fields and streamlines, bar charts, scatterplots in combination with graphical marks and visual encoding parameters. These abstract data elements can be embedded in concrete virtual environments, such as a virtual basketball court or airplanes.

Interaction & User Experience — Various studies and experiments were conducted on interaction and user experience in immersive volume visualizations. Some of them also satisfy our criteria for abstract visualization elements. For instance, Teylingen et al. [377] presented a tool for virtual heterogeneous data exploration and analysis. The internal data is hierarchically organized in customizable classes. Therefore, abstract data can be the basis for such a class. The system was demonstrated by means of visualizing molecular dynamic simulations of biochemical structures as well as the fluid dynamic simulation of a tilting rotor blade in hovering mode. In the context of this review, it is interesting that vector glyphs are used to depict the velocity field near the tip of the rotor. Here – already in 1997 – different glyphs and menus provided various interaction methods. Similarly concerned with interaction modalities, Hyde et al. [168] discussed an approach that offers a number of features for viewing and interacting with geological models in VR using the Oculus Rift. It offers human-centric navigation and manipulation, implicit surface editing, and visual conditioning. Volumetric grid data, including cross-sections, can be visualized and explored, and uncertainty data can be mapped to abstract and geological surfaces, e.g., in the context of drill hole planning.

2.4.1.7 Flow

Our analysis revealed four papers concerned with flow visualizations in a broader sense. While two demonstrate visualization approaches for flow data, the remaining papers focus on the design and evaluation of interaction methods and user experience by means of flow visualizations.

Trajectories & Movement — In an application paper, Hurter et al. [166] introduced FiberClay, an immersive multidimensional visualization system to visualize and analyze huge amounts of 3D trajectories in VR (see Figure 2.9). They demonstrated the applicability and usefulness of their approach by means of use cases and expert evaluations from the domains of air traffic control and neurology. Similarly, Homps et al. [158] present an approach for the interactive analysis of 3D trajectories in immersive environments but set the focus on the comparison of different selection modes in which different basic 3D shapes are deployed.

User Perception — Barrie et al. [16] focused on the evaluation of user performance when working with flow visualizations in VREs. They presented a study on the impact of immersion on the users' ability to analyze particle flows in a virtual environment. They concluded that an increased field of regard and a high degree of immersion can lead to better comprehension scores for the interpretation of particle flows.

Interaction — Other papers focused more on interaction techniques and methods associated with flow visualizations. For instance, Kageyama et al. [181] presented software for the visualization of 3D vector fields in a CAVE VR environment. They used the tracking of position and direction of stereo glasses and the wand controller to support interaction and updating of viewpoints. More recently, Mota et al. [255] developed the 3De lens, a focus+context visualization technique of multi-geometry data in VR. In their approach, they merged two categories of lenses - 3D and Cecal - to enable seamless analytical exploration of multi-geometry data using the focus+context paradigm in VR. As application cases, the aerodynamics of wind turbines were visualized with flow lines, and for the exploration of the

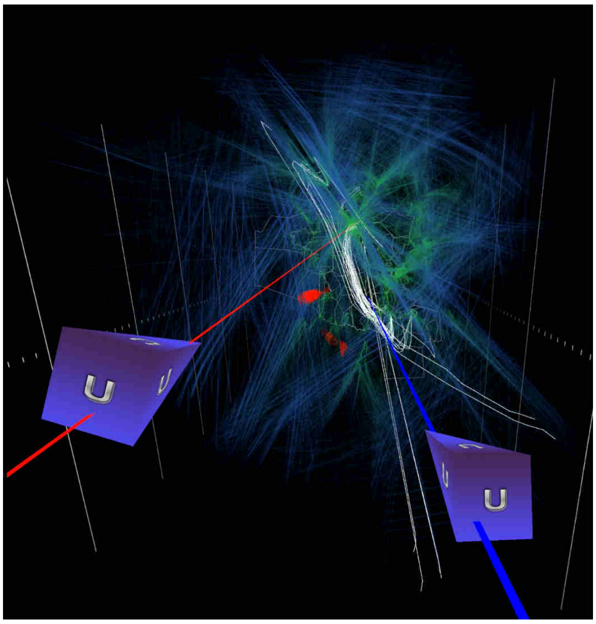

Figure 2.9: 3D trajectories inspected with Fiberclay [166] in an immersive VR environment. While the trajectories can have an inherent spatial meaning as, e.g., in flight trajectories of planes, the approach can also be deployed on non-geographic data, e.g., from the medical domain. Image courtesy of Christophe Hurter.

aneurysm, the lens patch provided depth information to improve the perception of surface shape and vessel-blood flow relations.

2.4.1.8 Other

In this category, we grouped papers that apply rare visualization techniques that could not be assigned to any of the other groups. Wijayasekara et al. [405] aimed to improve the usability of self-organizing maps (SOMs) for multidimensional data by providing an interactive neuron map visualization of SOMs as a 3D cube. They stated that the interactive 3D visualization helps to gain insight into the topology and relationships in the data. The visualization toolkit DXR of Sciat et al. [341] allows the visualization designer to create all kinds of custom visualizations, such as 3D bar charts, and is not limited to a certain set of visualization techniques. Schroeder et al. [332] visualized data as a combination of a bubble chart and a bee-swarm plot. Kraus et al. [198] conducted a study on 3D heightmap visualizations for comparative analysis tasks and compared them with juxtapositioned 2D heatmaps. Their results indicate a potential benefit of immersive environments for certain comparative tasks, such as estimating the relative offset of given locations in heatmaps.

2.4.2 Analysis Task

In the previous section, we have seen that the number of studies and their reported successes strongly depend on the visualization method considered. In this section, we shift our focus to analysis tasks and explore how different tasks were investigated in immersive analytics solutions. In a bottom-up approach, we identified seven types of analysis tasks and assigned each paper to one or more of them. Several publications present frameworks [74, 328, 341] or applications of visualization techniques [261, 323] without explicitly specifying concrete tasks and are, therefore, not regarded in this section. Dominating classes are ‘Visual Search’ and ‘Overview & Details on Demand’. Figure 2.3 gives an overview of the classification of tasks used along with the cited papers. In the following, we take a close look at each class of analysis task and summarize our findings.

Table 2.3: Literature review. Analysis tasks.

Analysis Task	References
Clustering / Classification	[6, 48, 116, 208, 266, 301]
Anomaly Detection	[6, 116, 266, 403]
Pattern Analysis	[143, 260, 301, 403]
Visual Search	[6, 18, 75, 96, 98, 152, 158, 213, 221, 266, 301, 332, 360, 364, 377, 389, 398, 399, 403, 422, 428]
Overview & Details on Demand	[15, 16, 19, 121, 149, 152, 158, 168, 181, 221, 236, 255, 332, 364, 377, 405, 421, 428]
Comparative Analysis	[166, 198, 199, 227, 236, 422]
Data Enrichment	[301, 311, 349]
Not specified	[26, 47, 71, 73, 74, 149, 209, 261, 265, 311, 323, 328, 341]

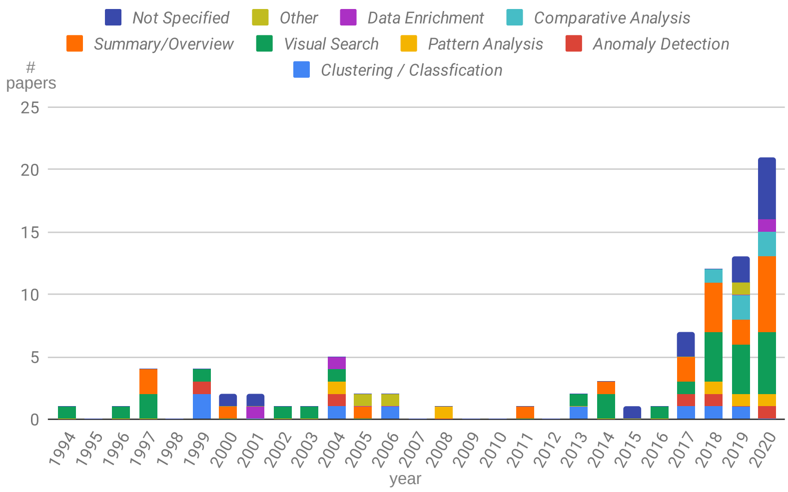

Figure 2.10: Literature review. Distribution of analysis tasks over time.

2.4.2.1 Clustering / Classification

Most papers that deploy clustering or classification tasks use them as a tool to compare differences in perception and analysis efficiency between different media. There are many works that make use of clustering or classfication as typical visual analysis tasks to evaluate different types of CAVES [6, 266], to conduct cross-comparisons between multiple media (2D screen, HMD VR, CAVE) [111, 301], or to compare HMD VR environments with 2D screen environments [116, 208]. For instance, Etemadpour et al. [111] compared the CAVE environment to a conventional 2D screen setup and measured the task performance of users completing various clustering tasks, such as counting clusters, finding the cluster closest to a given cluster, and detecting the densest cluster. We identified one paper that presents a cluster identification task as a use case without quantitative evaluation [9], and one paper with a qualitative expert evaluation in which cluster identification is treated as a task that can be easily solved with the presented technique [48]. While all previously mentioned works make use of scatterplot visualizations, the latter is the only one that uses clustering or classification tasks on another type of visualization – namely, a 3D parallel coordinate plot.

2.4.2.2 Anomaly Detection

Even though anomaly detection is an essential component of data analysis, it is rarely investigated in the context of immersive analytics. In our set of papers, only two explicitly deploy an outlier detection task in their user studies [116, 301]. In addition, Arns et al. [6] did not specify the anomaly completely but let study participants search for 'outstanding characteristics' in statistical data.

2.4.2.3 Pattern Analysis

Similarly to anomaly detection, pattern analysis is a popular data analysis task. Raja et al. [301] made use of a task in which participants should determine a trend in given datasets visualized as scatterplots. Not restricted to trend analysis, Nagel et al. [260] presented a system that is optimized for the detection of non-linear correlations and relationships in static and dynamic data visualizations. They proposed different visualization and interaction techniques to improve this goal and presented use cases to demonstrate the applicability of their approach. Similarly, Günther et al. [143] presented a system for fast prototyping of immersive visualizations. As a showcase, they presented the analysis of flocking rules in agent swarms.

2.4.2.4 Visual Search

Especially in exploratory analysis scenarios visual search tasks are often deployed. Many papers describe visual search tasks in volume visualizations [96], geographic visualizations [152, 389], scatterplot visualizations [6, 266, 301, 364, 377], flow visualizations [422], and network visualizations [18, 75, 98, 213, 360, 398, 399]. While in most papers the task is described as a generic way to explore data in order to build new hypotheses, in several works specific tasks were described and deployed in user studies. For instance, path tracing [18, 75, 398, 399] and target finding [98, 360] tasks in network visualizations or defined target feature search in scatterplots [6] or space-time cubes [389].

2.4.2.5 Overview & Details on Demand

Giving an overview of the underlying data is often claimed in papers presenting immersive analytics visualization techniques and applications. Especially papers presenting or using geographic visualizations [15, 152, 168, 364] and flow visualizations [16, 181, 255] argue for great overview capabilities of the respective approach. Besides geo and flow visualizations, also other approaches make this claim. For example, Teylingen et al. [377] presented abstracted volume visualizations of molecules and argued for their capability to convey the structure of the molecules and give an overview of the explored data space. Similarly, Wijayasekara et al. [405] stated that visualizing SOM neurons in a 3D cube helps analysts to understand the topology of the network and to get an overview of relationships in the high-dimensional data.

Apart from overview capabilities on certain topological data types, Bellgardt et al. [19] explicitly elaborated on their details-on-demand approach in the visualization and exploration of high-dimensional data. They presented a visualization technique for the inspection of a single data point as an immersive landscape glyph. Moreover, Hayatpur et al. [149] try to exploit improved spatial memory capabilities by lay-outing a users' analytic provenance graph in virtual space. Results of their qualitative user study indicate beneficial effects of the provided spatial layout of the workflow for data exploration and data understanding.

2.4.2.6 Comparative Analysis

We identified six papers that describe comparative analysis tasks with different visualization objectives, such as protein structures [236], the comparison of flows on maps [422], and 3D heatmaps [198, 199]. The framework "FiberClay", tailored for comparative analysis and comparison procedures on trajectory sets, was presented by Hurter et al. [166]. In their approach, it is possible to compare sets of 3D trajectories with novel interaction techniques. More on a conceptual level, Liu et al. [227] evaluated different layout strategies for small multiple visualizations and deployed a comparative analysis task in which participants compared the length of bars in multiple 3D bar charts.

2.4.2.7 Data Enrichment

Label placement is a frequent and essential task in AR. The required dynamic positioning of text snippets brings together both the classic challenges of label placement in static 2D drawings and the issues arising from viewpoint movement, distance changes, and occlusion in 3D. However, label placement was rarely investigated in the context of abstract 3D visualization. One rare example is presented by Azuma and Furmanski [8] who evaluated label placement algorithms, including both cognitive and perceptual issues, and found no clear relation between their label movement metrics and the users' performance. However, they did find indicators that label overlap is a critical factor in readability, and therefore the choice of the right placement algorithm depends on the use case (much vs. less change in user viewpoint). In addition to label placement, two other works present techniques for data enrichment - in graph visualizations as interactive selection and manipulation of nodes [349] and in scatterplot visualizations in terms of point selections and custom annotations [301, 311].

2.4.3 Paper Types

In this section we group and analyze papers based on their paper type. As shown in Figure 2.11, papers of the type 'Evaluation' dominate the considered spectrum of papers and are relatively evenly distributed over time, whereas most papers of the type 'Technique" were presented in recent years. In the following, we take a close look at each class of paper type and summarize our findings.

2.4.3.1 Technique

In the analyzed corpus, we identified only a small number of technique papers. To be classified as a technique paper in the scope of this survey, the main technique presented should be a new visualization or interaction technique for immersive analytics. Many of the papers that we investigated deal with the transfer of existing visualization techniques into the immersive analytics domain and are therefore considered as application papers unless the transfer requires major design considerations and adaptions to the new environment.

Visualization Techniques — Several works present new techniques that make use of properties provided in the immersive environment. For instance, Bellgardt et al. [19] present a new technique for visualizing multi-dimensional datapoints as immersive glyphs. In their tool Gistualizer a single multi-dimensional datapoint is visualized by creating a scene around the observer based on datapoint properties. Similarly, Zenner et al. [428] transform abstract process models into virtual 3D environments, turning the exploration of complex process models into an interactive and multi-sensory VR experience. Making use of the huge space provided in virtual environments, Hayatpur et al. [149] try to exploit the

Table 2.4: Literature review. Paper types.

Paper Type	References
Technique	[19, 48, 71, 149, 198, 199, 255, 349, 354, 428]
Evaluation	[6, 16, 18, 75, 96, 98, 111, 116, 121, 208, 209, 213, 221, 227, 266, 301, 332, 389, 398, 399, 403, 421, 422]
System	[47, 73, 74, 143, 260, 261, 265, 328, 341, 405]
Model	[26]
Design Study / Application	[9, 15, 115, 152, 158, 166, 168, 181, 236, 311, 323, 360, 364, 377]

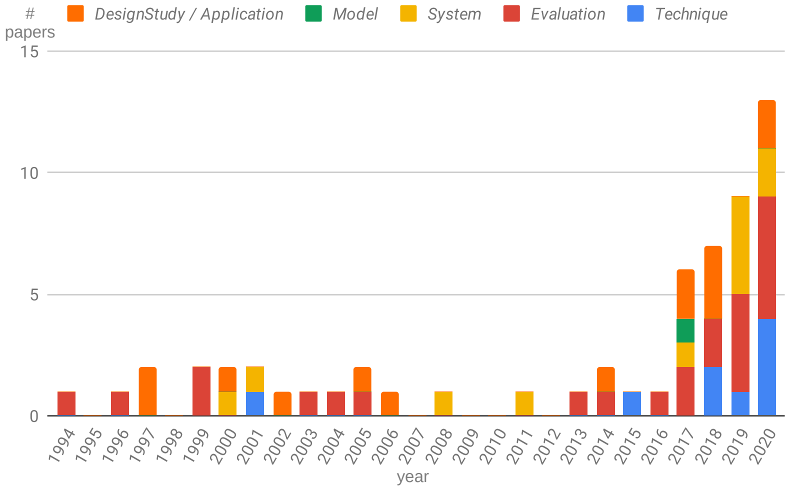

Figure 2.11: Literature review. Distribution of paper types over time.

"endless" workspace available in VR and present a technique for analytic provenance in which analysis steps are spatially lay-outed and can be explored by the user.
Other works adopt and extend or modify existing visualizations. For instance, Mota et al. [255] present 3De Interactive Lenses, a technique for focus&context flow visualizations of multi-geometry data and Kraus et al. [198] present a technique for the comparative analysis of 3D distributions in VR environments. In the latter approach, several 3D heatmaps are stacked on top of each other and can be shifted vertically into each other. Thereby, all heatmaps share a baseline, and local comparisons are eased.

Interaction Techniques — Various papers focus on interaction techniques that are only available in immersive environments and have a decisive influence on the exploration and interpretation of an abstract visualization. For instance, Slay et al. [349] present two egocentric interaction techniques within AR environments for selecting nodes in 3D graphs. Using fiducial markers, they deploy arm extension and ray-casting techniques to facilitate the selection of objects in 3D space. Similarly, bug focusing on navigation, Sommer et al. [354] present semi-immersive navigation techniques for "naturally" interacting with a stereoscopic 3D visualization. In their setup, they use a 3D screen in combination with shutter glasses and a stimulus pen as an input device for their interactions.
Butscher et al. [48] work in a hybrid visualization environment in which they visualize 3D PCPs in AR on top of a touch table. In their work the authors investigate novel interaction techniques tailored to the hybrid environment. In their framework (ART), it is possible to interact with the PCPs via gestures and touch interactions on the table. Also taking distance from pure XR environments, Cordeil et al. [71] use tangible "embodied axes" as a controller to interact with abstract 3D visualizations. Qualitative

expert feedback, as well as quantitative results of a controlled user study, indicate that their introduced interaction modality increases the accuracy in selection tasks, for instance, in scatterplots.

2.4.3.2 Evaluation

Evaluation publications considered in this survey can be divided into three high-level categories: (i) papers presenting the adoption of an existing visualization technique in immersive environments and evaluating its applicability for immersive analytics, (ii) papers that deal with the evaluation of fundamental human factors in visualization applications, and (iii) papers comparing the conventional medium screen with novel MR/AR/VR mediums for observing abstract visualizations.

Assessing the Suitability of AR/VR — In the considered corpus, six papers are concerned with the evaluation of a certain abstract visualization when observed in an immersive environment, with the focus being on the assessment of user performance, usability, etc. of the immersive visualization technique itself as well as on a comparison of different variants of the visualization or the medium used for presentation. Barrie et al. [16] used a cave system for animated data visualizations of particle flows. In their evaluation, they compared different configurations of the CAVE system for the given visualization task and measured the impact of the number of walls used and the presence of stereopsis. Similarly, Cordeil et al. [75] compared two different VR mediums in collaborative network exploration tasks. The evaluation comprised a direct comparison of a CAVE system with an HMD setup on different network exploration tasks, such as shortest path finding or triangle counting. Other works focus on the assessment of certain visualization techniques and compare their variants in immersive environments. With the focus on map visualizations, Yang et al. [423] evaluated different map and globe representations for origin-destination flow visualizations in VR. They compared four different visualization variants (3D exocentric globe, flat map, egocentric globe, curved map) and assessed their performance in rudimentary tasks, such as distance or direction comparisons of flows. Based on the results of this evaluation, they conducted a follow-up study [421] in which they investigated different visual encodings for flow maps, focusing on the presentation of flow lines (straight, curved, height variance). Working on scatterplots, Fonnet et al. [121] evaluated different axis positioning and layout strategies for scatterplots in immersive environments. They measured how well participants could estimate relative point locations and reconstruct the exact values (x,y,z) of datapoints with different axis layouts. Similarly, Drogemuller et al. [98] evaluated different navigation techniques for search tasks in immersive graph visualizations.

Human Factors — The second category of evaluation papers is concerned with the evaluation of fundamental human factors inherent to immersive environments, with the focus being on the assessment of the impact of immersion on the observation of abstract visualizations. Early on, Ware et al. [398] compared the performance of users when inspecting 3D networks in 2D, in stereo 3D, and in VR. In their follow-up work [399], the authors elaborated their evaluation and conducted an exhaustive user study, comparing 2D with 3D mediums for inspecting 3D networks and investigating the impact of depth cues for data understanding in 3D network visualizations.

Focusing on visual perception, Krekhov et al. [209] made use of the property of stereoscopic vision in VR for highlighting. In the technique "deadeye" they propose, highlighting is achieved by displaying highlighted objects on one eye only, making their appearance more dominant to the observer. Similarly, Whitlock et al. [403] compared time and accuracy of information conveyed over five different visual channels when observed on a conventional screen, in AR, or in VR.

Others focus on the assessment of human factors in certain visual analytics tasks. For instance, Kraus

et al. [208] evaluated different stages of immersion and their impact on cluster identification tasks in scatterplots. In their evaluation, 2D and 3D scatterplots are displayed on screens and in VREs, and user performance for cluster identification tasks is assessed. As one finding, they identified disadvantages when being fully immersed within a visualization due to overview issues and a limited reach of the field of view. Yang et al. [421] take a closer look at this particular problem and compare two techniques (zooming and overview+detail) for maintaining an overview while navigating within abstract scatterplot visualizations.
Besides studies on immersion and stereoscopic vision, also other evaluations of human factors fall into this category when being concerned with the assessment of different conditions only in AR/VR without comparison to monitor screens. For instance, the work by Liu et al. [227] compares different layout strategies for small multiple visualizations in VR space and evaluate how well participants get along with different layouts in various tasks, or the work by Lee et al. [221] compares different designs for collaborative environments to each other.

Comparing Media — The third category of evaluation papers includes evaluations that compare the conventional medium screen with immersive VR/AR mediums for inspecting abstract visualizations. The visualization most often evaluated in this sense, is the scatterplot. For example, many works compare conventional monitor screens to immersive environments (e.g., Cave [6, 111], Cave + HMD [301], HMD [96, 116]). Thereby, measuring user performance in typical scatterplot analysis tasks is a popular choice, such as distance estimation, cluster segregation, and outlier detection.
Besides scatterplots, graph visualizations are also quite popular for comparing different media. For instance, Belcher et al. [18] compared the medium screen with an AR HMD for path tracing performance on graph visualizations. The focus was on evaluating the benefit of stereo-vision for path tracing in 3D graphs, which were displayed as 2D projections on a screen, in 3D on a screen, and in 3D in AR. Their results indicate that 3D outperforms 2D in terms of performance, usability, and understandability of the graph structure. However, they could not identify any benefits for using AR, as participants performed equally well in the screen 3D condition. Similarly, Kwon et al. [213] evaluated user performance for graph visualizations observed with different mediums (screen + VR HMD), but focused on different visualization variants, graph sizes, and task types. They found that participants were faster, used fewer interactions, and gave more correct answers for large graphs when in VR.
Of course, inter-media comparisons are not restricted to scatterplots and graphs, and there are various comparisons that make use of different visualizations. For instance, Wagner et al. [389] evaluated the implementation of a space-time cube in a virtual environment (HMD) and compared it with a monitor screen setup. Their qualitative evaluation included human factors such as usability, required learning curve, mental workload, and simulator sickness. Even rarer techniques were used Schroeder et al. [332] who deployed bubble charts and bee-swarm plots to investigate differences in user perception between AR HMDs and monitor screens.

2.4.3.3 System

We identified several system papers that present platforms for the development of abstract visualizations in immersive environments. Most of them present their platforms within a specific visualization demonstration or use case but describe the extensibility and broad applicability of their system.

Research Prototypes for CAVEs — Sawant et al. [328] provide an overview of a whole collection of visualization systems for CAVE VR environments and propose the "Tele-Immersive Data Explorer", a system with a distributed architecture for collaborative, interactive visualizations that includes a combination of interactive desks and a CAVE VRE. Similarly, Nagel et al. [260] presented such a system for creating dynamic scatterplot visualizations with sound cues that offer a list of audiovisual tools. Their system builds on the modular system architecture of a previously developed approach [261] with a similar scope. Wijayasekara et al. [405] proposed “CAVE-SOM”, a framework system that allows the visualization of self-organizing maps as 3D cubes in CAVE environments.

Authoring Toolkits for HMDs — A number of systems are available that focus on rapid visualization prototyping for HMD-based VR environments. Most of them make use of the Unity Gaming Engine [371]. For instance, Cordeil et al. [74] focus on the provision of an authoring toolkit (ImAxes) for axis-based visualizations, such as parallel coordinate plots and 2D/3D scatterplots. Similarly, Sicat et al. [341] presented DXR, a visualization authoring system that allows the interactive creation of different visualization types such as scatterplots, bar charts, geo-visualizations, and flow visualizations. There are several other systems with similar scope, but partly for different environments, for example, AITK [73] for VR, and MRAT [265] for MR. Besides Unity-based frameworks, other approaches have emerged that aim at VR deployment on native Java VMs [143] (scenery) or on the web [47] (VRIA).

2.4.3.4 Model

In the considered corpus of papers, we identified only one paper as a model paper. Billow et al. [26] reflected on how a system can be evaluated in the domain of immersive analytics and presented a heuristic for evaluating immersive analytic systems. Their assessment used ten points to evaluate immersive analytic systems.

2.4.3.5 Design Study / Application

The second most common paper type in our corpus comprises papers with applications of known techniques and approaches that have not previously been used in immersive environments or in different constellations. For instance, Azzag et al. [9] demonstrated the usage of VR for the interactive exploration of multimedia databases.

In the domain of bioinformatics, several works investigate the applicability of existing visualizations and analysis approaches in immersive space, for instance, by experimenting with the display and interactive exploration of genome visualizations [115, 360] or protein networks [236]. Similarly, many geo applications try to exploit benefits of the three-dimensionality provided by VR/AR. Examples are volumetric data and 3D surface visualizations [168, 364], space-time-cube visualizations of time-dependent geo-trajectories [323], or sensor data visualizations in geo-context [15].

Similarly, for flow visualizations different applications investigated different directions. For example, they deal with the display and interactive inspection of fluid dynamics [377], vector fields [181], or 3D trajectories [158, 166]. Another example is presented by Reipschläger et al. [311] that demonstrates the creativity in application papers that goes far beyond a simple adaption of existing screen-based techniques in immersive environnments. They combined powerwall displays and AR headsets. In their approach, 2D visualizations presented on large powerwall displays are extended and connected to each other with AR visuals.

2.4.4 Technology

In the context of abstract data visualization, we identified three different categories based on technology: monitor, CAVE, and HMDs. Figure 2.12 shows the distribution of papers over time with regard to the Technology category. Although this data sample is relatively small, it reflects historical developments of immersive technologies and their research applications. Around the year 2000, CAVE-related papers were very popular. Although the initial CAVE had been invented some years earlier by Cruz-Neira et al. [77], research institutions started around the year 2000 to acquire CAVEs and use them in various research projects. Interestingly, only a few papers are from around the year 2010, indicating that VR was a rather unpopular topic at that time. It is not surprising that with their commercial success and increasing affordability, HMDs were used extensively for various research projects in the last few years. In the following, we take a close look at each type of technology and summarize our findings.

Table 2.5: Literature review. Technologies.

Technology	References
Monitor	[9, 311, 354, 398, 399]
CAVE	[6, 16, 19, 75, 111, 152, 181, 260, 261, 266, 301, 328, 360, 364, 405]
HMD	[15, 18, 26, 47, 48, 71, 73, 74, 75, 96, 98, 115, 116, 121, 143, 149, 158, 166, 168, 198, 199, 208, 209, 213, 221, 227, 236, 255, 265, 311, 323, 332, 341, 349, 377, 389, 403, 421, 422, 428]

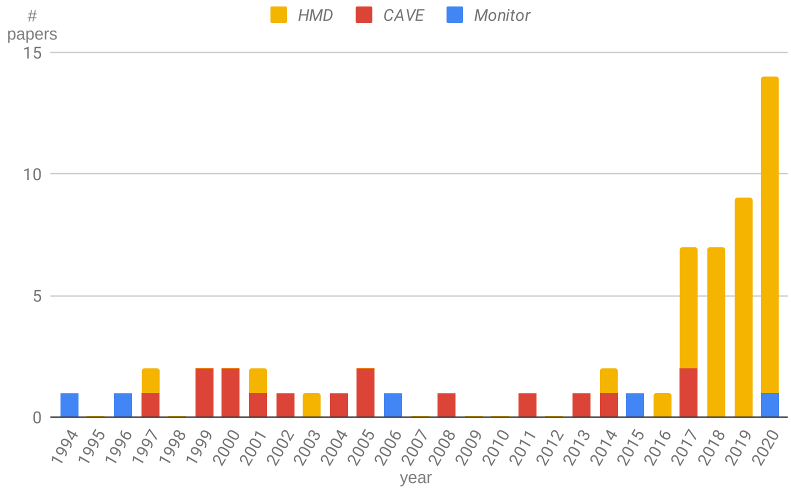

Figure 2.12: Literature review. Distribution of technologies over time.

2.4.4.1 Monitor

We identified few papers that make use of monitors to create immersive analytics environments for abstract 3D visualizations that meet our requirements (Section 2.3). Various approaches use a 3D monitor in combination with shutter glasses [9, 181, 354, 398]. For instance, early experiments were made with high resolution/frame rate monitors in combination with stereo glasses and head tracking in the context of a path tracing task [398, 399] or the perception of visual variables, such as shape, color, and texture, representing multimedia data in scatterplots [9]. Recently, [354] used a commercial approach, the zSpace – a passive stereoscopic monitor supporting spatial tracking of the head and a specific pen – to explore abstract variables of a car model and a biological cell. Moreover, this work used hybrid-dimensional visualization, using a 2D monitor to visualize a simple 2D network representation and the zSpace to explore the data semi-immersively. We also found papers that experiment with large powerwall setups. For instance, Maes et al. [236] used a powerwall-setup and compared it to HMDs, while Reipschläger et al. [311] investigated the interplay between powerwall and AR headsets in an evaluation study of augmentation approaches for static 2D visualizations.

2.4.4.2 CAVE

For more than two decades, CAVEs were popular devices in VR-related research with a broad range of applications. However, due to high acquisition and maintenance costs, accessibility is limited to a relatively small circle of researchers and end-users.

Hardware Diversity — A wide range of different CAVE setups exists - also in the context of abstract visualization. Some papers provide a very detailed definition of the CAVE setup used. To give an example of an early, well described CAVE configuration: Symanzik et al. [364] used a CAVE of 12x12x9 ft. where stereo images were projected on three walls and the floor and shutter glasses were used in combination with a magnetic-based tracker, a cyberglove, and a handheld wand. This 'wand' is a handheld input device for interaction with 3D objects and menus in the immersive environment and is a popular device for analytical tasks in the CAVE [181, 405]. Most CAVE setups use passive stereoscopic glasses but a few, especially older approaches, use shutter glasses [16, 260, 364]. While some works only describe modules of their framework in most detail, others limit themselves to a very shallow description of the deployed hardware. For instance, Raja et al. [301] elaborated on the value of head tracking in a CAVE during a study and describe this technological component in most detail while not elaborating too much on the composition of the CAVE itself, and Ferey et al. [115] describe a CAVE-like setup with two rear-projected orthogonal screens without further elaboration.

Domains and Tasks — Especially since the commercial success of HMDs, collaboration has become a central CAVE domain. [328] presented a collaborative environment with an interplay between CAVE and interactive, non-immersive desks. Various analysis tasks and techniques were explored in the CAVE context, including network analysis [360], particle flow analysis [16], data mining and statistics [260, 266], scatterplots [111], as well as glyphs [19]. Cordeil et al. [75] compared CAVE and HMD for collaborative network analysis.

2.4.4.3 HMD

HMDs were used in the aforementioned comparative studies in which monitor and projector-based visualization were compared with immersive visualizations [116, 236]. HMDs became very popular and were used in aforementioned application, technique, and evaluation papers due to their high affordability – especially in contrast to the previously-discussed CAVEs. In this survey, we avoid a direct comparison of different hardware setups under consideration of their suitability for abstract 3D visualization due to the enormous landscape of different devices and a very narrow field of quantitatively assessed setups, which makes it difficult to objectively judge and generalize the contextual quality of a certain device. For comparing hardware specifications of state-of-the-art AR/VR HMDs, we refer to up-to-date online resources (e.g., [20, 315, 407]). The following paragraphs break down which HMDs were most dominantly deployed for certain abstract 3D visualizations discussed in this paper.

Early HMDs — HMDs were already used in early works. For example, Teylingen et al. [377] developed the Virtual Data Visualizer system by using a SGI Indigo Elan, the Crimson Reality Engine, and a standard HMD setup. Slay et al. [349] used the DSTO InVision system to evaluate different interaction modes in AR. Belcher et al. [18] used a SONY Glasstron LDI-100B HMD in combination with an ELMO mini camera as well as a cardboard disc with tracking markers to visualize manipulable virtual elements.

VR-HMDs — Nowadays, a number of HMDs for consumer and/or industry usage are on the market. Checa et al. [59] evaluated the usage of HMDs in the context of immersive serious VR games and came to the conclusion that the HTC VIVE and Oculus Rift seem to be the most popular VR-HMDs. This view seems to be also supported by the selection of our papers: In work discussed in this paper, the HTC VIVE was used for parallel coordinate plots [74], graph visualizations [26, 98], interactive lenses [255], scatterplot layouting strategies [121], map and flow visualizations [422], comparative analysis [199], highlighting techniques [209], and cluster identification [208]. The Oculus Rift was used, among others, for collaborative scatterplot analyses [96], comparative 2D/3D graph studies [213], visualizing geological uncertainties [168], and space-time cube visualizations [389].

AR-HMDs — Although AR-HMDs are generally less often used for immersive analytics research on abstract 3D visualizations, the most popular AR-HMD is Microsoft Hololens. It was used in research projects in this context, e.g., for the representation of space-time cubes or scatterplots [71, 323]. Other papers contain a mix of technologies, such as Maes et al. [236], in which an HTC Vive, Oculus Rift, and a Power Wall setup is used. Besides native AR-HMDs, VR HMDs can be used as AR headsets if extended by see-through cameras attached to the headset. For instance, Butscher et al. [48] used a HTC VIVE VR-HMD with an Ovrvision Pro stereo see-through camera as an AR-HMD. In addition, there are papers that do not actually use the technology but suggest using HMDs in future projects, e.g., in [143]. In addition, various frameworks support the use of both AR and VR. For instance, the paper on the DXR framework demonstrates its applicability with both, the Hololens and the ACER VR headset [341].

2.4.5 Environment Type

In the following, we slightly shifted the focus from the technology used to the environment created by the technology: we clustered the papers along the Virtuality Continuum and distinguished between VR and AR applications. Both VR and AR technologies are applied in the context of Immersive Analytics. As Figure 2.6 indicates, VR has been the dominant environment throughout the considered period, but AR has seen an increased focus in the past few years. For differentiation, the Virtuality Continuum is often used [249]. In this section, we give some examples. However, since the previous section dealt with the technology being used to create immersive experiences, we will not discuss all works in depth in this section again and only reflect on the two environment types on a higher level.

AREs — We identified only a small number of examples that fit our constraints and make use of AR environments. Similar to Slay et al. [349], Belcher et al. [18] used AR through an HMD equipped with cameras (either stereo or mono). Fiducial markers in the real world were used to position virtual elements and track interaction devices. Saenz et al. [323] visualized a space-time cube in AR. Whitlock et al. [403] conducted a direct comparison between AR and VR (and screen), investigating differences in the perception of visual variables.

Table 2.6: Literature review. Environments.

Environment	References
AR	[47, 48, 71, 143, 265, 311, 323, 332, 341, 349, 354, 403]
VR	[6, 9, 15, 16, 19, 26, 47, 73, 74, 75, 96, 98, 111, 115, 116, 121, 143, 149, 152, 158, 166, 168, 198, 199, 208, 209, 213, 221, 227, 236, 255, 260, 261, 266, 301, 328, 341, 360, 364, 377, 389, 398, 399, 403, 405, 421, 428]

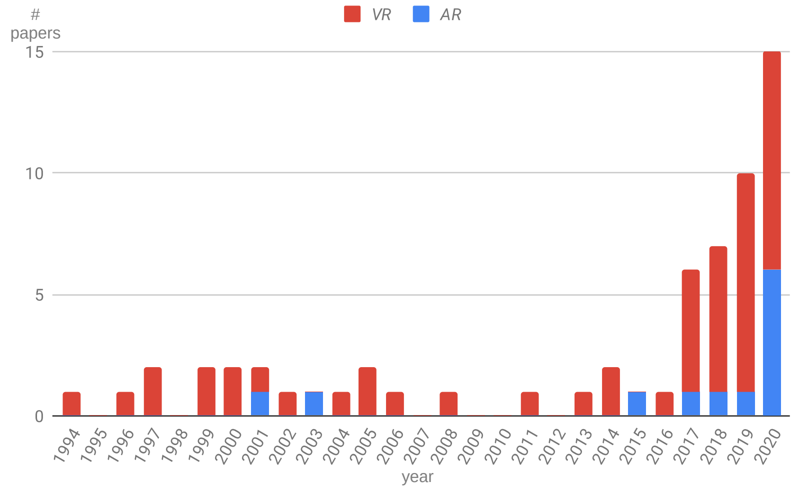

Figure 2.13: Literature review. Distribution of environment types over time.

***VREs* —** A number of approaches use classical multi-sided CAVEs in the context of VR [6, 16, 19, 181, 260, 261, 266, 328, 364, 405]. In other papers, software was developed with CAVEs in mind, without actually discussing their practical use [152, 360]. Approaches like that of Ferey et al. [115] use minimalistic CAVE setups and operate in VR. There are VR-related papers comparing CAVEs with other technologies, e.g. with 2D screens [111] or with HMDs [75, 301]. Many of the papers discussed here focus on HMDs as a technology [26, 74, 96, 98, 116, 166, 168, 198, 208, 209, 213, 236, 255, 341, 377, 389, 398]. As long as no additional camera is added to the HMD to project the Real World into the Virtual World, these are pure VR approaches. There are fewer approaches that use HMDs in the context of web-based visualization, e.g., Baltabayev et al. [15] used a phone/cardboard approach.

***Hybrid Environments* —** Some papers are on the borderline between AR and VR. For instance, Ware et al. [399] used head-coupled stereo viewing in combination with a monitor, Azzag et al. [9] used a stereoscopic monitor in combination with a data glove, and Sommer et al. [354] used a zSpace, which allows tracking of head and stylus when wearing passive stereo glasses. Butscher et al. [48] used a tabletop setup combined with a see-through HMD for collaborative analysis of multidimensional abstract data. Several frameworks support both AR and VR applications. For example, Gunther et al. [143] presented a Java-based VR/AR framework, Sicat et al. [341] the Unity-based DXR system for rapid prototyping, and Butcher et al. [47] a flexible web framework.

2.4.6 Data Types

Using this last dimension, each paper was analyzed with regard to the data considered in the respective visualization approach and grouped into one or more classes of data type. Most of the papers are concerned with the visualization of tabular data - i.e., independent data items with multiple dimensions listed in a table. Besides, we also found several papers on the visualization of network, field, and geometry data. In the following, we take a close look at each type of data and summarize our findings.

Table 2.7: Literature review. Data types.

Data Type	References
Tables	[6, 9, 15, 19, 47, 48, 71, 73, 74, 96, 111, 116, 121, 149, 158, 166, 199, 208, 209, 221, 227, 260, 261, 265, 266, 301, 311, 323, 328, 332, 341, 364, 377, 389, 403, 421, 428]
Networks	[18, 26, 75, 98, 115, 236, 349, 360, 398, 399, 422]
Fields	[16, 152, 181, 198, 341, 377, 405]
Geometry	[143, 168, 209, 236, 255, 354]

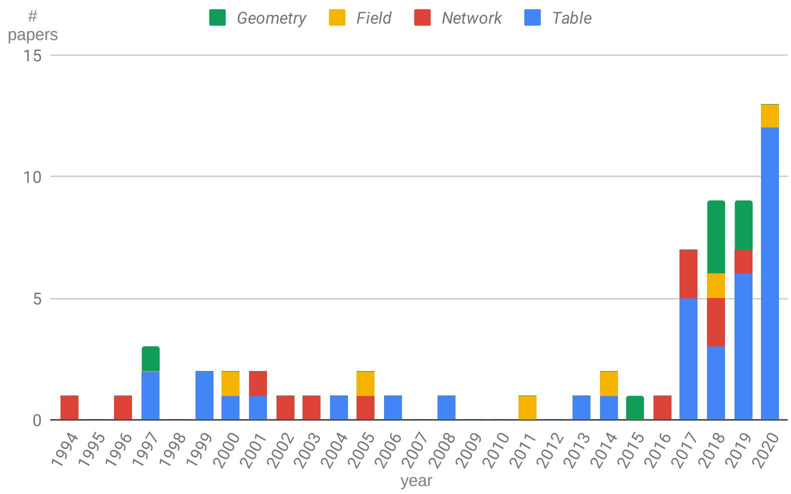

Figure 2.14: Literature review. Distribution of data types over time.

2.4.6.1 Tables

Table data, as classified by Munzner et al. [259] comprises multidimensional data where a data point is composed of a set of attributes. This type of data is the most frequently deployed in the considered corpus. In most cases, the papers do not specify the concrete meaning of the data used, but rather focus on the description of its properties and refer only to three-dimensional [121] or higher-dimensional data [48, 74, 96, 209, 261, 341]. In some works it is specified more precisely that the data are statistical data [6, 364], features from multimedia data [9], or higher-dimensional data that has been reduced to three dimensions [111, 116]. We also found two papers that explicitly stated that the underlying data was artificially created [208, 301]. Some papers describe table data with georeference [15, 328]. The data, therefore, also consists of data points with multiple attributes, but one of the attributes is a geocoordinate, which can be used to position the data on a geo-map. We found six works that make use of multivariate datasets in which each data point is defined for a series of time steps, i.e., for each tuple of data item and time, there are different attributes that compose a data entry in the table. Two of the six papers consider the temporal development of abstract sensor data [199, 260], while the other four deal with geospatial multivariate data [158, 166, 323, 389].

2.4.6.2 Networks

Network data includes data items that are interconnected. The relation between data entries is established with (weighted) references between data points. In several papers, the authors resort to artificially created network data to conduct controlled user studies with data sets with restricted properties [18, 398, 399]. Several papers do not further describe the origin and meaning of the datasets used, but rather describe their properties [26, 98, 360]. Additionally, we collected three papers that make use of biochemical data that describe the structure of genomes [115, 360] and proteins [236] as networks.

2.4.6.3 Fields

The class of field data contains all datasets in which values are associated with cells on a (2D/3D) grid. We found four papers which are concerned with vector field data [181, 198, 341, 377]. Helbig et al. [152] made use of climate data, which contains certain values (e.g., wind, temperature, precipitation) for each geolocation for a number of time steps. Similarly, Mota et al. [255] used time-variate surface information and explored flows in the data. We found one more paper that specifically focuses on flow data of particles [16]. Wijayasekara et al. [405] visualized the network structure of neurons in self-organizing maps (SOM). The underlying data can be classified as field data since each value is associated with a certain location on a 3D grid.

2.4.6.4 Geometry

Geometry data refers to datasets with data entries containing information about their spatial position. As we limited the scope of this survey to abstract 3D visualizations, only papers with abstract visualization elements remained in the final paper corpus. We found papers working with geometry data from the real world, such as plant [354] models. Others focus on microscopically small volume structures, such as redox-modified cysteines [236], or fly embryos [143]. In addition to volume geometry data, several papers also use data that establishes spatial positioning by means of stored geolocations, such as earth surface information [168, 255].

2.4.6.5 Reflection

Data types such as fields, real-world geometry, and data with georeferences lend themselves more naturally to a 3D visualization, and as a consequence, they also might allow creating simple yet intuitive interactions for navigation. Where 3D coordinates do not necessarily have a natural interpretation, such as for node-link network diagrams and some dimension-reducing projections of table data, more effort is required to conceive intuitive navigation. One associated challenge is to guide the user in the choice of insight-creating viewing perspectives.

While there is interest in using immersive technologies for the analysis of networks, the research is far from having explored a large portion of the corresponding design space. The freedom in the selection of the visualization idiom, encoding, interaction, and use of space is challenging, as the efficiency and effectiveness of different combinations are not yet well investigated and evaluated in immersive environments. In particular, for the large data sets from current applications, a big challenge is to create scalable approaches, e.g., by employing adaptive multi-level representations and abstractions.

2.5 Implications for Data Visualization

In this section, findings, lessons learned, and guidelines for the application of immersive environments in analysis tasks on abstract 3D visualizations are synthesized and summarized.

2.5.1 3D Structures & Depth Perception

Data that is visualized in 3D space (not necessarily spatial data) can profit from immersive analytics, for instance, through an improved depth perception of the analyst. The degree to which improved depth perception is beneficial for a certain analysis procedure depends mainly on the analysis task. For instance, results of quantitative user experiments revealed that distances between data points can be perceived more accurately in stereoscopic environments compared to monoscopic 2D displays [111, 116]. This fundamental finding is reflected in follow-up studies with more complex tasks like cluster identification [116, 208, 266], or outlier detection [116]. Hence, if, for a certain analysis task, distance estimation between two points in the 3D visualization is relevant, the deployment of immersive environments to observe the visualization may pose an advantage.

Frequently, the increased performance of participants working with stereoscopic settings is ascribed to improved depth perception. Researchers in various domains support the hypothesis that improved depth perception inherent to stereoscopic displays increase task performance in various spatial tasks [140] like measuring position and distance of data objects or path tracing in 3D graphs. Whitlock et al. [403] even consider that the improvements of depth perception inherent to stereoscopic viewing might alleviate the stigma of 3D visualizations. Wither et al. [412] present techniques to further boost depth perception with spatial cues. However, there are also critical voices in terms of adding additional depth cues to the visualization, especially for large data sizes with complex structures [230]. Another explanation for the improved performance of users in immersive environments could be the increased level of immersion of analysts in the data space. Arns et al. [6] came to the conclusion that the reason for the better performance of participants in the VR environment is the “true” three-dimensionality caused by the immersion of the user. This effect can also lead to a reduced learning curve in understanding more complex data structures [16], as shown for path tracing experiments in network graphs [18, 213, 399].

However, there is also research reporting different results. One drawback of AR was identified by Whitlock et al. [403]. In their study, participants had difficulties with decoding colors from visualizations due to the fact that virtual elements overlapped with the real-world environment. Therefore, they advise deploying color with care as a visual variable in AR environments.

In summary, research mainly reports positively about the use of 3D stereoscopic visualizations. However, designers have to be careful when working with low-resolution devices, large amounts of data, or additional depth cues since these can have a negative effect on the analysis result.

2.5.2 Navigation & Interaction

Immersive analytics opens new ways and possibilities for the design of interaction and navigation modalities. For instance, VR environments enable more intuitive and natural interactions (e. g., movement by walking, selection by grabbing). However, this great freedom of choice, in combination with the absence of guidelines and reference work for best practices, also leads to a high degree of uncertainty and arbitrariness when designing user interaction concepts for immersive applications.

Researchers report steep learning curves and poor performance of users with unfamiliar, direct interaction approaches.

Immersive analytics allows users to interact with 3D objects in 3D space directly and is, therefore, more engaging. Symanzik et al. presented scatterplots of statistical data in a CAVE environment and claimed that the visualization encompassing the user is “inviting interactions with the data” [364]. The more ‘natural’ and ‘intuitive’ interaction modalities associated with immersive environments are often cited as the reason for improved accuracy [6, 18, 349]. However, results regarding task completion times differ. While some could show that task performance increased with the new interaction techniques, others found the opposite effect of higher task completion times with immersive interactions (e. g., [6]).

Besides interactions for manipulating visualizations, various sources report improved navigation capabilities in immersive environments. For instance, Kwon et al. [213] found that navigation in 3D graphs in VR was more manageable compared to navigation in 2D graphs in screen-based environments. Immersive environments in which head movements control the perceived view on a visualization allow intuitive control of the viewport and improve navigation in 3D space. Raja et al. [301] concluded from a user study comparing the performance of users in various tasks on scatterplots when working in VR and in screen-based environments that “head tracking showed a strong trend in favor of its use”. Task completion times, disorientation, and usefulness ratings of users, as well as personal observations, led to their conclusion about the usefulness of head-tracking. Similarly, Hurter et al. [166] found a general benefit in the intuitive control of a user’s viewport induced by head movements. However, without some orientation support people in VR environments might suffer from motion sickness or disorientation [279].

Various sources report on difficulties inherent to immersive interaction modalities. While Wagner et al. [389] found many aspects in favor of using immersive environments for the interactive analysis of space-time cubes, such as higher usability scores, higher user preference, and lower workload, users performed slightly worse when immersed. The authors attributed this finding to users’ unfamiliarity with VR and the resulting interactions. Arns et al. [6] conducted a user study comparing the interaction capabilities of participants in a VR environment and in a screen-based environment. Even though they found that participants needed much more time to complete the given cluster selection task when immersed, they also found significant differences when taking the users’ experience with VR into account. Users who were more familiar with VR were much faster compared to novice users. Hence, they concluded that interaction difficulties and the associated decrease in efficiency could be due to a steep learning curve and lack of familiarity with novel VR environments. Similarly, the line of argumentation of other works is that unfamiliarity is a major obstacle that makes usability comparisons of novel immersive and familiar screen-based environments difficult [261].

In summary, research reports positively about the integration of 3D interaction and navigation when analyzing data in VR or AR settings. However, there seems to be a steep learning curve in 3D navigation for users, which negatively affects completion time. Establishing some common grounds or guidelines for 3d user interaction and navigation might mitigate this negative effect.

2.5.3 Hardware

Immersive environments can be created with different mediums. The choice of the medium can have an impact on the effectiveness and efficiency of visualizations perceived within the therewith created AR or VR environment. For example, current AR HMDs, such as the Microsoft Hololens, have a very limited field of view, which affects the impression of immersion [422]. This can be circumvented by creating the AR environment with see-through VR [349], in which cameras capture and manipulate the real environment and display it in a VR HMD.

Augmented and virtual reality environments have both benefits and drawbacks compared to each other. AR environments provide better contextual awareness and reduce the likelihood of simulator sickness [349], while VR environments maximize immersion and enable remote collaboration in a completely shared environment [96]. In addition to differences in perception and usability, the hardware also differs in manageability and costs. While a CAVE setup is bulky and expensive [364], head-mounted solutions are much cheaper and more common [96].

Moreover, general technical limitations of state-of-the-art AR and VR technologies must be taken into account. For example, when discussing the poor performance and high task completion times of users when comparing their immersive environment with a conventional screen-based setup, Belcher et al. [18] refer to hampering properties of their deployed device, such as low resolution, limited field of view, and color and contrast characteristics.

With regard to virtual reality hardware, we have seen a transition from CAVEs to HMDs VREs. While CAVE setups were very popular before the mid-2010s, the technology was displaced by more mobile and cheaper HMD solutions. The reason for that could be the broad offer of different, consumer-ready VR HMDs by various manufacturers. With that trend of VR being used by a wider range of people, we can expect immersive analytics to become more broadly applicable and easier accessible in the future.

In summary, there is a trend from CAVE environments to more flexible, affordable, and mobile HMD devices. This transition will pave the way for an increase in immersive analytics applications. However, designers have to carefully consider the application domain and weighing up the benefits of 3D against the drawbacks like lower resolutions or worse color characteristics in comparison to typical 2D screen setups.

2.5.4 Guidelines & Common Practice

Especially in earlier years of immersive analytics, researchers reported difficulties in designing user interfaces and visualizations for immersive environments. The visualization space in immersive environments is large and, in contrast to conventional screen-based environments, not restricted to a certain area (i. e., the screen). This complicates the design of visualization frameworks. For instance, Symazik et al. [364] discussed where a geo-map visualization could be optimally placed in the virtual environment. The design of user interfaces and menus in immersive environments is similarly difficult, and the adoption of conventional screen-optimized menus is not always feasible [377]. Similarly, several sources report on difficulties when designing interactions for their visualizations due to the absence of guidelines in the field [266, 301]. According to Whitlock et al. [403], this also holds true for most basic research. The authors state that we still lack empirical grounding for how to best visualize data in immersive environments. In their work, the authors try to counter the issue by initial studies on visual variables, comparing the effectiveness and expressiveness of different variables like size, color, orientation, and depth in scatterplot visualizations. While comparing AR, VR, and screen, their results indicate differences between all three media. Even though this gives us a first glance on medium-specific differences of the effectiveness of visual variables, exhaustive guidelines for visual variables, gestalt laws, or pre-attentive perception as they are available for traditional monitor screens are still outstanding.

In summary, designing VR or AR applications for data analysis is still a challenging task. Due to the vast design space of immersive environments and the lack of empirical research, guidelines are still rare. More research is needed to establish a common basis for future designers to rely on.

2.5.5 Collaboration

The use of immersive environments can offer several advantages for collaborative analysis tasks on abstract 3D visualizations. For instance, Butscher et al. [48] discussed the potential of deploying AR for the collaborative analysis of multidimensional data visualized as PCPs. In their approach, the analysts are in the same physical environment and share the same digital content, allowing natural communication and coordination between collaborators. Similarly, Cordeil et al. [75] investigated the performance of users in co-located collaborative tasks on graph visualizations and compared a CAVE setup to an HMD VR environment. Their results suggest that both compared VR platforms perform equally well in most aspects for the tasks investigated. This and the fact that CAVE VREs are much more expensive, require more maintenance and are not available to the general public speak in favor of using HMD VR devices for collaborative tasks.

In addition, immersive analytics enables the natural collaboration of remotely located collaborators. Different approaches for remote collaboration on abstract data visualizations were presented. Leading arguments for the application of VR are that sharing the same visual space leads to better collaboration in visual data exploration tasks [96], improves communication [168], and makes collaboration more convenient due to direct interaction capabilities in shared visualizations [328]. Somewhat more rarely, nevertheless represented, is research on co-located collaboration in VR environments. For instance, Lee et al. [221] compared different designs for collaborative co-located VR environments and argue for the highly flexible design of the shared workspace as an advantage of VR.

In summary, immersive environments support collaboration because of more natural interaction between users and a shared visual space for data exploration. These findings are independent of the underlying hardware favoring HMDs since they are less expensive and open to the general public.

2.6 Discussion and Open Research Areas

There are many papers describing new techniques for immersive visualizations, evaluations of existing non-immersive approaches deployed in AR or VR, comparisons between different immersive and non-immersive media, immersive visualization systems, and applications of immersive environments for abstract 3D visualizations. However, there are hardly any taxonomy and model papers that focus on the application of abstract 3D visualizations in immersive environments. More and more research deals with the assessment of differences between mediums on abstract 3D visualizations and the identification of potentially beneficial properties of immersive environments in restricted settings. However, there are few generalizable guidelines and recommendations as to when and where the use of immersive environments can bring benefits. Initial observations in various studies suggested that even established visualization paradigms could be overwritten in immersive environments. For instance, gestalt laws or the order of visual variables according to their effectiveness, may be perceived differently in such environments, which could change the way visualizations should be designed for IA applications.

Therefore, more fundamental research is required to address general issues of immersive visualization and to provide general guidelines for the application of visualizations in AR/VR.

Over the last decades, a large number of different immersive technologies have been evaluated as media for displaying abstract 3D visualizations and compared to conventional, non-immersive analysis environments. While research uniformly points to advantages in immersive environments, such as direct manipulation capabilities of visualization elements bypassing indirect input modalities (e. g., mouse/keyboard interactions), interaction difficulties are often cited as a hindering factor for efficient analysis procedures. High degrees of freedom and interaction constraints (e. g., text input, coding) complicate various user interactions. The constant progress in technology leads to the continuous development of new interaction modalities for immersive environments, which have to be evaluated individually. Moreover, technological advances in immersive devices could also affect the effectiveness of certain visualizations and overwrite evaluation results of previous studies with outdated technologies. As previous research suggests, the level of perceived immersion is decisively influenced by factors like multisensory stimulation, display resolution, and fidelity/photo-realisticness of the virtual environment. Increased levels of immersion can, in turn, influence visual analysis tasks. For instance, if the user is allowed to touch, feel, or even smell data points with haptic VR gloves or HMD extensions, the illusion of actually dealing with real objects is enhanced.

Another popular justification for poorly functioning VR/AR scenarios is that immersive environments and accompanied input modalities are highly unfamiliar to most participants. Therefore, VR/AR environments might already increase their effectiveness if users are better trained and more familiar with the new environments. However, this could have a decreasing effect on other dimensions such as excitement and engagement, which could be increased in new AR/VR environments just by the fact of low levels of familiarization. Novel interaction paradigms invite further assessments. For instance, virtual teleportation is a popular technique to compensate for the limited physical space in VREs and needs to be carefully evaluated in contrast to physical walking or other alternatives such as VR treadmills or redirected walking.

In short, more research is needed to assess the actual impact of technology differences (resolution, fidelity, multisensory stimulation) and user familiarization on user performance in immersive visual analysis tasks. Further, studies on outdated devices and technologies may need to be repeated on newer devices that lead to higher levels of immersion. Of course, the results of previous studies can be used as a starting point for formulating new hypotheses.

The deployment of immersive environments for data analysis is largely independent of data types. The usefulness and applicability of immersive visualizations depend on the target analysis task and the chosen visualization type. We observed that none of the investigated papers contained abstract 3D visualizations for *Text* data. We assume that the main obstacle factors for this are missing or incorrect input modalities for text in VR/AR and non-optimal technical constraints of immersive devices, such as low resolution, which make reading text in the respective immersive environments difficult. Nevertheless, the great potential of plain text analysis in immersive environments should be considered carefully in future research.

While most IA papers focus on fundamental research on common visualizations (e. g., scatterplots, node-link graphs, geo-map visualizations), only a few make use of or present 3D adaptions of rare visualization techniques such as dense pixel visualizations, sankey diagrams, chord diagrams, arc diagrams, cartograms, stream charts, dendrograms, or complex glyph visualizations. While it is important to evaluate the basic properties of immersion and their impact on visualization efficiency and effectiveness in combination with new interaction and visualization design conditions, the assessment of more complex niche visualizations deployed in immersive environments would be highly interesting.

There is no general and uniform framework, library, or programming language that can be used to generate visualizations for immersive applications quickly. Certainly, there are different frameworks that allow quick prototyping of a certain set of visualizations, but much effort is needed to create the above-mentioned types of visualizations. In addition, existing IA authoring toolkit papers frequently point out the difficulty of completing all steps in the reference model of visualization [395] in order to create a visualization from scratch in immersive environments due to restricted code/text interaction capabilities. Therefore, the three main steps of applying data transformations, visual mappings, and view transformations are mainly restricted to the latter two, and hardly any framework supports data transformation procedures in immersive environments. In this regard, establishing more standards for data handling and transmission might be a fruitful research direction and help for future designers.

Several 3D visualizations have proven to work poorly on 2D monitor screens. However, influencing factors induced by immersive environments could balance out or even eliminate some of the main disadvantages of such visualizations. Therefore, it might be reasonable to consider re-evaluating visualizations with a bad reputation, such as 3D bar charts in immersive environments. In some cases, the optimal approach might be a combination of 2D and 3D visualizations, allowing smooth transitions or links between them and taking advantage of both types.

Future research should not only focus on fundamental research but also explore the application of immersive technology for more advanced types of visualizations or the differences in collaboration when exposed to small restricted rooms in comparison to open space environments. Furthermore, there is a need for uniform development and authoring environments that facilitate the process of creating new types of device-independent immersive visualizations and make immersive visualization accessible to non-experts without advanced programming skills.

In previous research, abstract 3D visualization was mainly used in exploratory and confirmatory analysis procedures. However, we see a great potential of immersive visualizations for information presentation scenarios where the only goal is to convey information to an observer [302]. Previous research has shown that immersive environments can enhance memorability, increase engagement, and even intensify emotions. Such factors could help to make information more accessible, understandable, and lasting. These properties could also prove helpful with regard to gamification and gameful learning. The application of immersive environments is also studied in various other research domains, such as in educational research (teaching scenarios with pupils), psychology (phobia treatment), and entertainment (game development). Future research in the field of immersive analytics should tie in with research in other domains, apply cross-domain knowledge transfer, and use findings and insights from other domains as a basis for new hypotheses.

Basic research in immersive analytics could reveal certain potentially beneficial properties in immersive environments, such as improved spatial memory, direct object manipulation, natural navigation, and improved depth perception through stereoscopic vision. However, most results are very task- and condition-specific and cannot be generalized. Depending on the task and condition, the visualization expert must assess whether properties that have proven useful in other cases apply to the current problem and must modify the visualization or environment to take advantage of potential benefits. Future research should try to establish general guidelines for the design of immersive visualizations to support the optimization of immersive analyses.

Although IA research pointed out various benefits of deploying immersive environments for the analysis of abstract data, AR/VR devices are not yet established media that are widely used in the industry for the visualization of abstract 3D data. The main obstacles for this could be (a) the lack of established end-user visualization environments (such as Tableau and others for non-immersive visualizations) for creating and exploring visualizations in VR, (b) high efforts to present immersive visualizations to a large audience, and (c) usability constraints, such as uncomfortable and tedious head-mounted displays or bulky and expensive setups.

2.7 Conclusion

We conducted a survey on publications of immersive analytics approaches for abstract data visualization. The publication selection was based on a keyword search and manual inspection. The base set was expanded by scanning the references of matching papers, resulting in a corpus of 48 papers covering a period of 27 years. A key observation from our survey is a surge in the number of publications in recent years. While this is not clear evidence that immersive environments are already accepted for abstract data analysis after years of skepticism, it does show that the design space and potential are being explored in current research projects. Furthermore, we can see that a variety of aspects is being investigated regarding data type, visualization technique, and paper categories. However, while CAVEs played a central role in the early years of VR-related research, research on environments based on VR HMDs clearly dominates today. This may be due to the relatively inexpensive devices, the easy setup of such an environment, which is almost plug-and-play, and the broad support by available software for content creation. In addition, the controllers of current HMDs allow for quite intuitive interaction that goes beyond the standard desktop setup.

Despite the diversity of research topics covered in the publications investigated, there seems to be no structured exploration of the design space. As the results from studies are often quite specific to the conditions and tasks used, better characterization and specification would help to enable replication, but also a more structured approach to evaluating the potential of IA for abstract data visualization. Similarly, there is no common code base, such as a toolkit or framework, that supports fast prototyping of general solutions, and much effort is put into developing necessary basics for each of the projects. However, in the course of our survey, we have discussed a number of toolkits that already implement a wide selection of visualizations discussed here (e.g. [73, 265, 341]). Although there are prototypes, they are not widely used in the community, and developers tend to start their projects from scratch. Thus, we can see the potential for a community effort to create supporting toolkits that can be used for prototyping. Further initiatives are needed to develop common standards as a basis for general IA toolkits optimized for visualizing abstract data.

It is interesting to note that although the number of research projects using immersive technologies is increasing dramatically, the amount of abstract data visualization in this domain is relatively small. This can be seen in the relatively small number of papers found based on our search criteria. Therefore, this area has much potential for new findings.

3

Foundations for Deploying Virtual Reality for Data Visualization

Contents

Previous work has shown in individual demonstrations that for some tasks and visualizations, the deployment of immersive environments can offer benefits - even for abstract data visualizations. In this chapter, the basic arguments for and against the use of VR for data visualization are discussed at a more general level. To demonstrate the first strategy to assess the applicability of VR for data visualization, three examples are presented that justify or motivate the deployment of VR. In the first part, a general approach is presented for how extending the visual context – also with the help of VR – can contribute to analysis verification and uncertainty analysis. Although the presented approach is generally applicable, even outside the field of immersive analytics, we demonstrate its applicability with VR examples. This part is based on our IVAPP paper [206], which focuses mainly on non-abstract visualizations and applications. The second part shifts the focus to the – in theory – generic applicability of VR for all data visualization and visual analytics approaches. Each step in a generic knowledge discovery pipeline is examined considering the use of VR instead of a conventional monitor screen setup as a visual interface. This work was published as a Springer book article [203]. In the third part, an exemplary basis for designing immersive analytics VR environments is evaluated: the adequacy of replicating real elements and properties in virtual environments. This final part was presented at the VisGuides workshop at IEEEVis [207] and critically discusses the many facets of replication that must be considered when designing immersive analysis environments.

3.1 Breaking the Curse of Visual Data Exploration: Improving Analyses by Building Bridges between Data World and Real World

Visual data exploration is a useful means to extract relevant information from large sets of data. The visual analytics pipeline processes data recorded from the real world to extract knowledge. Subsequently, the resulting knowledge is associated with the real world and applied to it. However, the considered data for the analysis is usually only a small fraction of the actual real-world data and lacks above all in context information. It can easily happen that crucial context information is disregarded, leading to false conclusions about the real world. Therefore, conclusions and reasoning based on the analysis of this data pertain to the world represented by the data, and may not be valid for the real world. The purpose of this chapter is to raise awareness of this discrepancy between the data world and the real world which has a high impact on the validity of analysis results in the real world. We propose two strategies which help to identify and remove specific differences between the data world and the real world. The usefulness and applicability of our strategies are demonstrated via several use cases.

3.1.1 Introduction

Nowadays, large amounts of data are generated and collected within mere seconds. As a result, constantly increasing amounts of information are available and subject to an increasing number of analytical data acquisitions and new technological possibilities with regard to gathering, storing and distributing data. Many people are interested in gaining knowledge from this data, for example, by using data mining algorithms or visual analytics [182] methods. Afterwards, the generated knowledge is applied on the real world where the used data originate from. However, there is a flaw inherent in our everyday analytical reasoning: The collected data is no perfect representation of the real world. Many facets of our surroundings cannot be measured with the necessary precision, if at all. Also, likely there exist factors influencing the analysis that we are not yet aware of and therefore do not measure. Since performing an analysis on incomplete and noisy data cannot lead to fully complete and correct results, we claim that *data is always wrong* to some extent. Consequently, the analysis might not generate valid real-world knowledge, but instead knowledge that is valid in the world represented by the data. For example, when measuring the speed of cars in a rally race, the collected data is necessarily rounded to a certain degree. Also additional factors, such as vertical accelerations might be neglected. Therefore, results of the analysis based on this data, such as the maximum speed or the average acceleration of cars, only deliver answers to the abstracted data on the rally race, not the rally race itself.

We draw attention to this important problem to which we further refer to as *the curse of visual data exploration*. Without a doubt, for some domains and tasks, the considered data can be sufficient to lead to similar results as if the entire real world would have been taken into account for the respective analysis. Still, each divergence in the data from the real world leads to a slightly less optimal result, and it is often hard to tell how much the data diverges from the real world. This issue has already been recognized and been part of researcher discussions all around the world, e.g., in the panel discussion of the 2017 IEEE Symposium on Visualization in Data Science (VDS). The related topic of uncertainty analysis is mainly concerned with the data gathering process and the validation of gathered data. Often, however, the problem does not lay in the data itself (e.g., faulty or missing data), but in the scope of the measured data (e.g., parameters not considered for analysis), which is usually not addressed by uncertainty analysis. To raise awareness and foster discussion, we examine the *curse of visual data exploration* (section 3.1.3) and provide possible strategies to *break the curse* (section 3.1.4) by focusing on projecting data and analysis results back into a more comprehensive real-world context. If the real world is not sufficiently described by our data, we are able to identify this by inspecting the resulting visualizations in the overall context of the real world and evaluate if all necessary data is considered in the analysis. We elaborate on the general usefulness of our proposed strategies and provide several examples of projects (section 3.1.5) in which we applied our proposed recommendations, even though they cannot yet be applied to every analytical use case (section 3.1.6).

3.1.2 Related Work

To identify the *curse of visual data exploration*, we studied published Visual Analytics pipelines in the literature and recognized a missing connection between the generated knowledge and the real world. Through all stages, different sources modify the data in a way that the data no longer fully corresponds to the real world. We first discuss related work in uncertainty analysis (section 3.1.2.1) followed by an overview about data validation (section 3.1.2.2). We position our work within the aforementioned works in section 3.1.2.3.

3.1.2.1 Uncertainty

As uncertainty occurs in almost every field of research, one important challenge is to find a generalized definition of uncertainty that can be applied to various domains. MacEachren et al. [232] recognized early on that uncertainty is a complex concept which needs to be subdivided into different components. Subsequently, suitable methods for the representation and processing of uncertainty are needed. Skeels et al. [344] surveyed the state-of-the-art and introduced a model identifying components of uncertainty in various fields of research such as information visualization. Their model divides uncertainty into three levels with increasing abstraction. *Measurement precision*, *completeness*, and *inferences*. Measurement precision deals with imprecise measurements of sensors which could be identified by confidence intervals. One level above is the completeness which describes the loss of information by using projections or sampling techniques. The highest abstraction reflects the inferences. Inferences describe the uncertainty of predictions of future values based on current data. Here, one challenge is that prediction models cannot predict a value if the underlying data has no similar data points. All levels are covered by an additional component (*Credibility*) which describes the trustfulness of the data source as well as potential disagreements when comparing among other sources. Gershon et al. [131] describes the challenges of visualizations in an *imperfect world*. In their taxonomy, they illustrate and summarize the different challenges that arise when gathering data from the real world. The resulting taxonomy is divided into two parts, on the one hand, the imperfection (uncertainty) during data acquisition. On the other hand, falsely represented data, for example, an overplotted visualizations or the use of an inappropriate device.

3.1.2.2 Data Validation in Visual Analytics

Several definitions of data validation exist in different domains, e.g., in the Unece glossary in statistical data editing [7] it is described as an action to verify if a value matches an allowed set of values. Also, Wills and Roecker [411] describe data validation as the ability of a model to detect variance in the data including, for example, the recognition of missing values and outliers. Over the years, a variety of outlier detection techniques such as anomaly detection [57], noise detection [309] or novelty detection in time series [88] have been proposed. Hodge and Austin discuss several outlier detection techniques in their survey [156] and identify three types of how outliers can be found based on different knowledge about the data. In the first type, there is no knowledge about which data points belong to the outliers. Consequently, outlier detection is based on a statistical distribution to classify whether a point belongs to a specific distribution or not. For the second type, each data point requires a label to indicate whether it is an outlier or not. Subsequent, classifiers are trained by using these labels to predict if a new data point is an outlier or not. This involves the generation of a classifier for detecting normal data points and abnormal data points. Type three again uses pre-labeled data like in type two with the difference that these classifiers trained only on the data points which do not belong to the outlier class. This can be used to determine whether a new data point belongs to the set of valid data points based on the training dataset. However, data points which do not belong to this set are not necessarily outliers. This approach for type three is similar to algorithms for semi-supervised recognition or detection tasks.

3.1.2.3 Positioning of our Work

Approaches handling data manipulation, for example, in a visual analytics workflow are mainly concerned about stages between data collection and knowledge generation. Our model differentiates from the current state-of-the-art by introducing a new validation step enabling the validation of whether extracted knowledge applies to the real world.

3.1.3 The Curse of Visual Data Exploration

"*The cost of bad data is the illusion of knowledge*" [370]. At the beginning of the data analysis process it is important to consider the quality of the collected data. Research nowadays is mainly concerned with improving the results of an analysis, both regarding performance and quality. Unfortunately, the quality of the data used for this analysis is often not ensured to be adequate for a given task. If the quality of the data lacks in detail, is faulty or incomplete, conclusions drawn from the analysis might only be referable to the data but not to the actual real world the data was taken from. Therefore, generated knowledge would probably not apply to the investigated research question as intended.

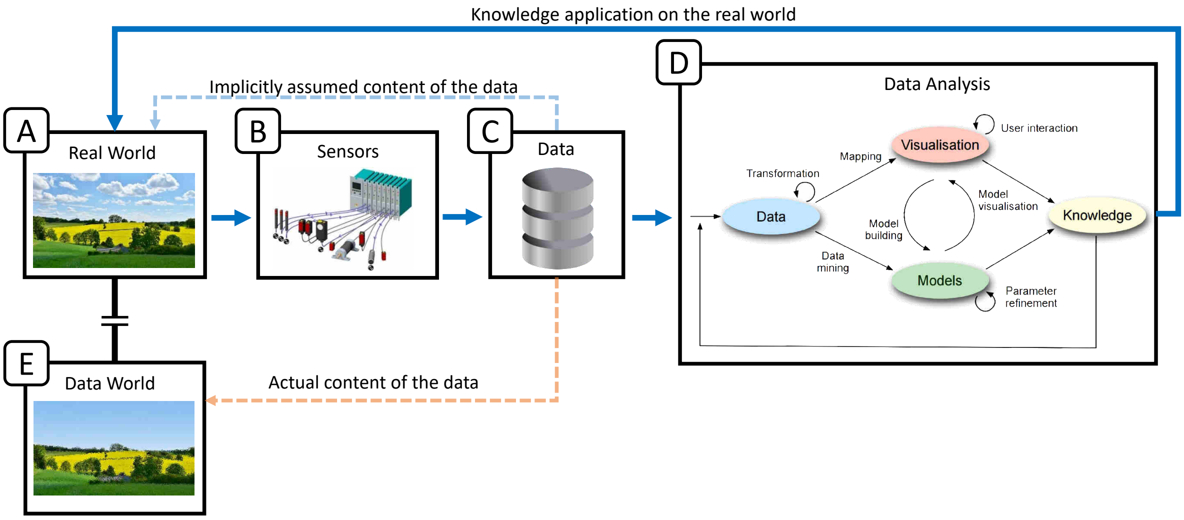

Figure 3.1: The curse of visual data exploration displayed in an extended visual analytics model. In common data analytics tasks, the procedure starts in the real world (A) where information is collected using sensor technology (B) and stored (C). Afterward, the gathered data is analyzed as described in various proposed visual analytics models. This process is here depicted in (D) at the example of the Visual Analytics pipeline by Keim et al. [190]. The knowledge generated by these models is often assumed to be correct in the real world implying that the gathered data represents a complete and correct copy of the real world. However, as the gathered data only contains a subset of aspects (the data world, E) that can have an influence on the analysis process, the generated knowledge may not be complete or even invalid in the real world.

A broader framework for common analytical workflows such as visual analytics can be seen in Figure 3.1. The collection of data is the starting point where data is obtained from the real world (Figure 3.1 A) using, for example, sensor technology (Figure 3.1 B). We refer to the real world as the world we want to analyze. Usually, this is the physical world around us, but it could also be a conceptual world like a stock market. Sensors capture properties and discretize continuous signals to digitalize real-world information (e.g., video recordings). There are various types of sensors, for instance, thermometers, pressure sensors, and cameras. Gathered data is merged and stored digitally as discrete values, abstracted to bits and bytes (Figure 3.1 C). Stored data is typically the starting point for common analytical workflows such as visual analytics [394]. In Figure 3.1 D, we inserted the Visual Data-Exploration Pipeline by Keim et al. [182] as an example for arbitrary visual analytics pipelines. Any other analytical workflow might be inserted here as long as the following two conditions are met: (1) They start with digital data as basis for the analysis and (2) their goal is to generate knowledge about the real world.

Finally, the output of the analysis (knowledge) is attributed back to the real world from which the data was extracted from. The generated knowledge is naturally assumed to be valid in the real world since all the input data came from the real world. In Figure 3.1, this assumption is depicted by a blue arrow going from the extracted knowledge to the real world. The fact that knowledge is valid in the data world does not mean that it is not valid in the real world as well. Some information from the real world is more important to the analysis than other, and usually, most of the real world information is not relevant for a given analysis task. Therefore, the validity of the knowledge in the real world strongly depends on how accurate and complete all important information sources have been measured. In some cases, analysts might be aware of missing aspects in the data that are hard or even impossible to capture. However, in general, the real world is a complex construct that is impossible to capture completely and correctly. Due to this fact, analysts have to assume that the data world and the real world are similar enough to transfer generated knowledge to the real world. This is what we call the *curse of visual data exploration*.

In detail, we refer to the *curse of visual data exploration* as the natural condition of incomplete or faulty data as a basis for analytical workflows, leading to a wrong association of generated knowledge to the real world. This association of knowledge would only be legitimate if the gathered data would completely, correctly and exclusively represent the real world. However, knowledge would also be transferable, if the data used for the analysis contains all information influencing generated knowledge. I.e., in practice it would be sufficient if all analysis relevant information would have been considered throughout the analysis. Factors that do not influence the analysis results (irrelevant real-world data) can be neglected. Whenever data is collected, there is a high chance that some important information is neglected that would impact the analysis process and therefore the generated knowledge. This loss of important real-world information can occur in several ways. Sensors may produce systematic or random errors, sample insufficiently or create somewhat unwanted biases in the data. Besides, the analyst may not be aware of factors that are not yet considered in the analysis (missing sensors). In statistics, such factors are also referred to as *lurking variables* [33]. Faulty data could also be introduced through abstracting procedures during the gathering process (e.g., aggregating, binning, sampling, digitalizing, discretizing). These complications throughout the gathering process lead to a discrepancy between the real world and the collected data. The world described by the data is, therefore, not perfectly representative of the real world (faulty, incomplete). In the following, we refer to the entirety of the gathered data as the *data world* (Figure 3.1 E). The analysis is conducted in the data world and, therefore, it is only ensured that generated knowledge is valid in the data world.

3.1.4 Breaking the Curse

In this section, we propose two strategies that aim at minimizing the chance to be affected by the *curse of visual data exploration*. The goal of these strategies is to make sure that the generated knowledge is not only valid in the data world but also in the real world. Both strategies follow the same principle of going back to the real world to validate the data or the results.

To break the curse, we aim to minimize or remove the gap between the real world and the data world. Since the analysis results are valid for the data world, they would also be valid in the real world if both are equal to each other. More precisely, it would be enough if both worlds were equal concerning all information that affects the analysis as the results would then be the same. In the following, we refer to this information as *analysis relevant information*.

Since the data world cannot realistically be an exact copy of the real world, the data world usually contains a subset of the information available in the real world. To ensure a valid analysis procedure and to allow inference of results to the real world scenario, the following conditions have to be true:

S1 Information contained in the data world is correct, i.e., it is not contradicting to real-world data.

S2 Dimensions (attributes) contained in the data world are also present in the real world, i.e., the data world is a subset of real world.

S3 Dimensions contained in the data world cover all the analysis relevant information of the real world.

In general, it is impossible to guarantee that all conditions (*S1*-*S3*) are fulfilled, e.g., due to imperfect measurement accuracy or the deployment of derived dimensions (artificial dimensions that do not directly reflect properties of the real world). However, it is desirable to optimize *S1*, *S2* and *S3* as much as possible.

S1 ensures that the collected data is correct and therefore not negatively impacting the analysis. Common causes for violations against *S1* are random or systematic errors in measurement devices. *S1* can be ensured by comparing each value present in the data world with its corresponding value in the real world. This procedure can be time-consuming if the data set is sufficiently large and sometimes even impossible if the measurement cannot be repeated. *S2* ensures that no additional data exists in the data world that does not describe the real world. This could happen if the dataset is a composition of multiple sources of which some are valid sources describing the real world and others are not. Validating *S2* can be done by checking if every individual dimension of the data world can be found in the real world. *S3* ensures that no analysis relevant data is missing. If analysis relevant data would be missing, the analysis can end up at faulty results as crucial information would have been neglected for the examination of the real world. Usually, the examined real world is complex, hampering the validation process and making automatic validation impossible. Therefore, user involvement is required. An analyst can apply their real-world knowledge to identify dimensions that likely influence the current analysis task. For example, if the task is to predict crop growth, the analyst would identify dimensions like solar radiation and precipitation as important. However, it is challenging to recall all possible variables influencing a process, especially if the analyst is not reminded of their existence in some way. This makes *S3* the hardest of the statements to verify. In the following, we propose two exemplary strategies to verify *S1*-*S3* with the aid of visualization. Currently, our strategies are limited to suited data and use-cases. For example, very abstract data such as multivariate results of questionnaires might not be optimal for the presented approaches as they have no spatial, temporal, volume or similar context that could be visualized easily.

3.1.4.1 Reconstructing the Real World

Since validating *S1*-*S3* can be rather complex and time-consuming, we propose a strategy for how their validity can be checked using visualization. During the analysis step, complex information is often abstracted and visualized making it easier to comprehend. We propose a similar strategy to check the validity of *S1*-*S3*. While we usually collect data from the real world and create data representations, it is also possible to go the other way around and use the data contained in the data world to reconstruct a subset of the real world.

If *S1*-*S3* hold true, this recreation should contain each aspect of the real world that is thought to be relevant for the analysis process. Comparing the visual representation of the reconstruction to the real world can help to reveal differences between the two, such as missing or faulty properties. The visual representation should aid the user in spotting differences which would be hard to notice by just looking at the abstracted data. With the aid of this reconstruction, the user can make use of knowledge about the real world to verify the validity of *S1*-*S3* by checking the reconstruction for inconsistencies or the absence of analysis relevant information. If there is anything in the reconstruction which is not present in the real world or differs from it, then either *S1* or *S2* is violated. Since visual representations help to understand a large amount of information quickly, this process is assumed to be a lot more efficient than validating every value in the data against its real-world counterpart as described earlier. Still, identifiable missing data can be of even higher interest.

For instance, sports analysts examining soccer matches are interested in analyzing regions their players can control (interaction spaces [356]). The orientation of players is an important factor in calculating these interaction spaces. Naturally, the space behind a player is smaller as the player has first to around in order to reach space behind him. This takes time and, therefore, decreases the area they can control behind them. Accordingly, when an analyst annotates the interaction space of a player manually while watching a video stream of a soccer match, the orientation of the players is subconsciously used in the analysts mental model. However, if in computer-assisted data gathering and analysis procedures, the data collected does not include the information about player orientation, *S3* is violated. Looking just at trajectories of players composed of single x- and y-coordinates, the player orientation could be assumed to be the direction of movement, and it is hard to realize that this attribute is actually missing. At this point, the reconstruction of the real world could improve the analysis process. In a three dimensional reconstruction based on the collected data, players can only be represented as abstract icons, as no information about their orientation and posture is available. The emerging visualization makes it impossible for the analyst to determine if players are running forward, backward, or sideways. Analysts would be able to realize that the inspected data is missing crucial information since their mental model is not in accordance with the visual representation. After identifying a dimension that is missing in the data world (in this example, player orientation), the respective data can be added to repeat the whole process until no more flaws are discovered.

3.1.4.2 Projecting Results back into the Real World

Our second proposed strategy does not aim to validate the data world directly but instead confirms analysis results by projecting them back into the real world. In analytical workflows such as visual analytics, generated knowledge is often presented via abstract visualizations like parallel coordinates [169] or glyph visualizations [125, 404]. These visualizations are useful to explain analysis results to humans, but they often include little context information about the real world. This creates a gap between the data world and the real world, even though the goal is usually to connect the generated knowledge to the real world. This separation makes it hard to judge whether the analysis results fit into the context of the real world. We argue that by projecting the analysis results into a space that is closer to the real world, users are enabled to reveal contradictions that would go unnoticed otherwise. Afterward, it must be ensured that the identified aspects are included in the data (*S3*) as well as measured correctly (*S1*). This proposed strategy has the additional advantage that problems within the analysis itself can also be spotted.

Figure 3.2: The GPS coordinates of tracked birds were projected into a virtual environment to analyze their behavior in a thermal spiral. From the eyes of a bird, you can see how other birds use the thermal spiral to move upwards in a circle. Also, the data is enriched by providing the surrounding landscape to recognize further factors which influence the behavior of birds. Using this visualization, it was possible to detect that specific winds, represented by cloud movement, could be responsible for a certain pattern in bird movement, which was previously not explainable.

Going back to the previously introduced soccer example, automatically determined interaction spaces of players are calculated based on players' speed, distance to the ball and running direction. Afterward, interaction spaces can be visualized as circles or circular sectors on an abstracted soccer pitch. However, a players movement direction does not necessarily reflect his or her body orientation. If the same visualization is projected into a video of the real world soccer match, it has reportedly been easier to spot that there is a problem with the used data for this analysis. In several expert studies performed in recent work [357], invited soccer analysts repeatedly reported that "*[...] they became more aware of a visualization's limitations and possibilities for improvement in the future. As, for example, soccer players were not represented by moving dots on an abstract pitch anymore but with the real persons, the experts noticed that the body pose is currently not always reflected correctly in the calculation of interaction spaces. If a player is running forwards or backwards, the resulting interaction spaces are identical. This exemplary problem could not attract attention outside of the video visualization as no data about the body poses are collected.*" (quoting [357])

3.1.5 Use Cases

To demonstrate the applicability of either reconstructing the real world (section 3.1.4.1) or projecting the results into the real world (section 3.1.4.2), we present a detailed use case for each of them. By the example of collective behavior analysis, we show how the data world can be reconstructed and verified. Afterward, we consider the use case of a criminal investigation, showing how the extracted analysis results can be projected back into the real world to verify the data basis for the analysis.

3.1.5.1 Collective Behavior

The first use case deals with the calculation of thermal spirals from tracked bird movement data. For this purpose, students from the University of Konstanz reconstructed a part of the real world based on the available data to validate if some features are missing as described in section 3.1.4.1. The movement data of the birds are provided by an online database called Movebank [406] managed by the Max Planck Institute for Ornithology [329]. Each bird is equipped with a GPS receiver to record the current position, direction and altitude. Researchers use this data to identify characteristics to see whether individual birds communicate to the swarm where thermal spirals are located. The integration of satellite images and elevation data into the virtual environment helps to investigate the external influences of thermal spirals better. During the analysis, the experts noticed that individual birds moved away a few meters from the swarm within a second. In the beginning, it could not be explained why the birds behave this way and how this influences the collective behavior. After integrating weather data as another data source into the virtual world, they noticed that a gust of wind has caught the birds and dragged them a few meters. This was only possible by representing the wind with the help of cloud movements. Seeing the clouds move in the same direction as the birds made it easier for the analyst to create this connection as compared to just looking at the numbers in the dataset. In Figure 3.2, you can see a part of the virtual environment out of the eyes of a bird. The current satellite image which matches the GPS position of the bird is located at the bottom so that users can always see the exact surroundings. This includes information like the height of the mountains as well as the land usages. The reason for displaying satellite images and elevation data is that thermal spirals behave differently, whether they occur over mountains or flat areas. The projection into a simulation enables experts to recognize missing features like the weather information.

3.1.5.2 Criminal Investigation

When investigating a crime scene, context information is undoubtedly crucial. Many side factors are overseen if only considering the fraction of the real-world data that is at first glance the most important data. For example, during a rampage in a city, the data available to law enforcement agents could consist of video material of surveillance cameras or pedestrians, mobile cell connections as well as email conversations of the suspects, reports of eyewitnesses, GPS locations of the suspects and much more. Still, it is impossible to consider all data available. Many dimensions of the real world that seem irrelevant would naturally have to be neglected to avoid processing overload (e.g., weather data, news data, traffic data or twitter data).

Algorithms might be able to extract features from video frames, analyze them and present condensed information to the analyst. This could, for instance, be achieved as follows: the algorithm searches in video frames for objects using deep neural networks and collects for each object all frames the object appears in. The result is a set of objects with corresponding trajectories of the objects. Subsequent visual analytic procedures could be deployed to analyze those trajectories. Knowledge deducted from this analysis is ascribed to the actual procedure of events throughout the incident.

Our strategy suggests projecting the extracted information back into the original footage. For instance, by marking matched objects in the video, or even by plotting the trajectories in a 3D reconstruction of the part of the city where the incident took place. This would create a geo-context that reflects the real world even more than video footage. Additional information such as weather or traffic information could be visualized as well. The original recordings could then be placed within this 3D world and be adjusted in time and space. Investigators could then walk through the actual crime scene, navigate in space and time and view recordings of interest.

Hereby, faulty or missing data might become apparent quickly. Analysts might detect objects in the video footage that were not detected by the neural network or classified wrongly. For instance, two objects with trajectories running along next to each other which are separated by a river in between could have wrongly be identified as a group by the algorithm. The additional geo-context given in the 3D reconstruction makes this error visible. The analyst learns that the algorithm did not consider geo-characteristics for the grouping of objects and is able to adapt the algorithm or at least consider its impact on the interpretation of remaining results.

3.1.6 Discussion

With our work, we want to rise awareness that the applicability of analysis results is highly dependent on the quality and completeness of the collected data. We discussed why collected data is not a perfect representation of the real world and introduced the concept of a data world in Figure 3.1. We elaborated that the discrepancy between the data world and the real world can affect the validity of analysis results and called this problem *the curse of visual data exploration*. Two strategies were introduced which can reduce the extent at which this curse occurs. We consider our proposed strategies as a step towards more sophisticated solutions to detect invalid or missing data measurements by using the concept of bringing back the collected data and generated knowledge into the real world. Of course, this concept has its limitations. Projecting data into a visual space that is closer to the real world is challenging as each scenario has to be handled application specific. Some data may not have a straightforward representation in the real world at all, especially if data describe a concept that is not visible in the real world. By allowing the user to incorporate real-world knowledge, this concept allows to detect data problems or missing data that would otherwise be hard to find.

We present two strategies in our work, one which reconstructs the real world from the collected data (section 3.1.4.1) and one which projects the generated knowledge back into the real world (section 3.1.4.2). While the strategies are similar to each other, they can lead to different results. Reconstructing the real world from data can highlight data dimensions which are not present in the real world as described by *S2* in section 3.1.4. For example, it could be that the data set has been manipulated by adding a dimension which does not exist in the real world. Recreating the real world from this data would result in a representation of this additional dimension which is visible in the recreation. Using the visual representation of the reconstruction as well as real-world knowledge, it might be easier to spot this additional data compared to just looking at the data set. Finding these problems with the other strategy is harder, as the projection of the analysis results into the real world would not contain this additional dimension anymore. On the other hand, the reconstruction strategy cannot identify problems introduced by the analysis concept which can be found using the projection strategy. Whether one of the strategies is superior to the other in specific cases is subject to future research.

In general, our introduced model can be applied to other domains which focus on the extraction of knowledge from data. For example, the Knowledge Discovery in Databases pipeline from Fayyad et al. [114]. In future work, we want to investigate how far back into the real world the knowledge should be projected to find most data quality problems. In the soccer example, one could go back to images, to videos or even to a reproduction of the match with real players. Our assumption is that going back this far is counterproductive as it could make it too complex to project the data into the real world.

3.1.7 Conclusion

To overcome the illusion of knowledge generated by invalid or incomplete data, we extend current visual analytic pipelines with a validation step to detect data measurement errors (errors in the data world - i.e., in the data that is considered in the analysis). Our concepts are based on the idea that the extracted knowledge is projected to a representation of the real world to test if the knowledge fits to the real world. We discussed this generic problem and two exemplary specific solutions. It is notable that they are not generic and applicable for any kind of data.

3.2 Accommodating Virtual Reality in the Visualization Pipeline

The previous section has exposed the discrepancy between the subject of analysis (real world) and the actual data on which the analysis is performed (data world) as a critical weak spot in visual analysis pipelines. In this section, we demonstrate how VR can help to verify the correspondence of both worlds in the context of InfoVis and VA. Immersion allows the analyst to dive into the data world and collate it to familiar real-world scenarios. If the data world lacks crucial dimensions, then these are also missing in created virtual environments, which may draw the analyst's attention to inconsistencies between the database and the subject of analysis. When situating VR in a generic visualization pipeline, we can confirm its basic equality compared to other media as well as possible benefits. To overcome the guarded stance of VR in InfoVis and VA, we present a structured analysis of arguments, exhibiting the circumstances that make VR a viable medium for visualizations. As a further contribution, we discuss how VR can aid in minimizing the gap between the data world and the real world and present a use case that demonstrates two solution approaches. Finally, we report on initial expert feedback attesting the applicability of our approach in a real-world scenario for crime scene investigation.

3.2.1 Introduction

Nowadays, data is collected at any time and in any context, in order to gain knowledge about real-world circumstances. Collecting data means storing digitized snapshots that reflect properties of the real world at a specific point in time. We use the term *data world* to refer to the sum of related snapshots collected for a particular use case. To eventually gain knowledge from the collected data, visual data exploration approaches, executed on the data world, are mostly indispensable due to the constantly growing amount of data. However, the snapshots contained in the data world never completely reflect reality, but rather a small fraction of the actual real world that lacks contextual information. Crucial context information can quickly be overlooked, which may lead to false conclusions about the real world. Hence, any analysis carried out on the basis of the data world provides results that are at most applicable to the data world itself and not to the real world as initially intended.

In previous section (3.1), we raised awareness of this *curse of visual data exploration* and proposed two potential strategies which help to identify and eliminate differences between the data world and the real world. The first strategy revolves around projecting analysis results back into the real world to confirm these results. Based on the projection of analysis results into a space which is closer to the real world, the user can consider additional contextual information to reveal contradictions that might otherwise stay unnoticed. The second strategy reconstructs the real world from the data contained in the data world to verify its validity. The reconstruction is intended to consider all relevant aspects of the real world for the analysis. This enables the user to compare the reconstruction with the real world to uncover differences such as missing features. For each of the two strategies, we provided a concise proof of concept.

Now, we build on this foundation and elaborate in detail how VR will help us to bridge the gap between the data world and the real world in the future. In the recent past, VR applications have already gained in importance, in particular in the computer graphics sector. For the information visualization and visual analytics domain, several approaches have been proposed as well [141, 294, 429], but the potential of virtual reality has not yet been fully explored and established. To show that virtual reality is helpful for problems that occur within visual data exploration, we must prove that virtual reality has a place in the established information visualization pipeline [62, 97, 146], too. Accordingly, in section 3.2.3, we introduce our concept of virtual reality for information visualization and emphasize that VR is not inferior to conventional media (printouts, screen-based displays) when applied correctly. Instead, we argue that it is possible to obtain visualizations in VR that are as suitable as those displayed using conventional media. Moreover, we show that specific properties inherent to VREs can be used to extend existing visualizations or even to establish new ones. After laying this foundation and defining the place of VR in the information visualization pipeline, we demonstrate in section 3.2.4 which advantages VR has to minimize the effects or even to solve the initially described curse of visual data exploration. Ultimately, in section 3.2.5, we describe an extensive real-world use case in crime scene investigation, applying both of the aforementioned strategies to break the curse. Overall, our contribution consists in (a) identifying dimensions to compare visualizations displayed on different media, (b) providing an overview of currently existing major drawbacks of VR for data visualization, (c) listing conceivable present and future possibilities of VR, and (d) discussing potential benefits associated with the use of VR to minimize the gap between the real world and the data world.

3.2.2 Thought Experiment: Hypothetical, Ideal VR

We investigate how and to what extent the discrepancy between the data world and the real world (see Figure 3.1 A, D) can be minimized by deploying VR and the first strategy (reconstructing the real world). As previously stated (see section 3.1), the gap would be minimal if the data world used for the analysis perfectly resembled the real world. Therefore, we need to develop a theoretical framework to discuss this important topic at a hypothetical level where the data world perfectly resembles the real world and is perceived in a VRE. By doing so, we first assume a VRE that perfectly resembles the real world *(1)* and elaborate on accompanied advantages *(2)*.

1. What if a VRE is actually perceived like reality? At least in theory, it is possible to see VR as an equivalent medium for perceiving visualizations. Consequently, we can imagine a hypothetical VR environment that perfectly reflects a real environment, for example, a so-called “Substitutional Reality” environment [342, 363] that replaces each real object with a virtual representation. In our case, instead of substituting each real object with an alternative virtual object with similar properties, it is replaced by a virtual representation of itself. For instance, a room with furniture should in our case in VR be represented exactly as it is in the real world. All objects are positioned where they are in the real world. As a result, anyone entering the “duplicated world” in this example would physically walk through the real room and perceive the virtual room (which looks like the real room). Any cloned, “virtual” object could be implicitly touched and interacted with as the virtual world can be considered a “duplicate” of the real environment and each virtual object has a real counterpart.

2. What would be the advantages? As soon as we have constructed a perfect representation of the real world, we need to illustrate why we should not simply make use of a conventional medium (such as a regular monitor screen) to inspect the replication of the real world, instead of a VRE. As a first step, we show in section 3.2.3 that VR has a place in the established information visualization pipeline. In section 3.2.3.1 we demonstrate how the ideal VR can be embedded in the information visualization pipeline, while in section 3.2.3.2 we show how perfect a reconstruction of the real world must be in order to be perceived and considered as equivalent by a user. We use the resulting findings in section 3.2.4 to verify whether VR is actually preferable to a conventional screen in our scenario.

3.2.3 VR in the Visualization Pipeline

In this section, we present a hypothetical, ideal VR as part of the previously presented thought experiment and demonstrate the equality between the hereby created VRE and a conventional screen-based setup with regard to the visualization pipeline (section 3.2.3.1). Furthermore, we discuss to what extent all prerequisites and assumptions made during the thought experiment are necessary to maintain equality between the compared media (section 3.2.3.2).

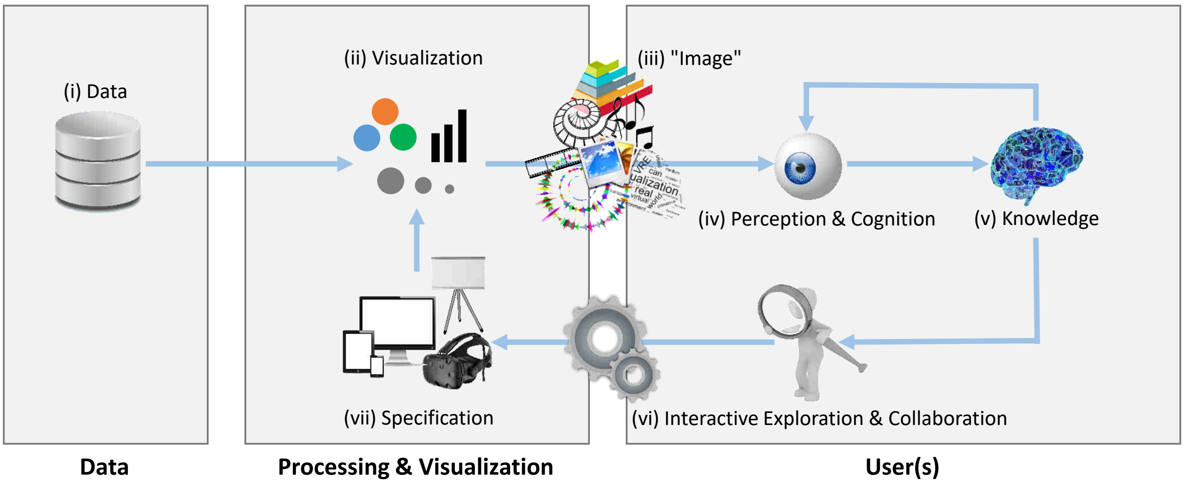

Figure 3.3: Visualization pipeline adopted from van Wijk et al. [380]. Data (i) is transformed and mapped (ii) based on specifications (vii). A resulting "Image" (iii) of the visualization is perceived (iv) by the analyst who can apply prior knowledge (v) to interpret the visualization and to acquire new knowledge (v). Subsequent interactive exploration (vi), possibly through collaboration (vi), may lead to the adaption (vi) of the specification (vii).

3.2.3.1 Embedding the Ideal VR in the InfoVis Pipeline

Since we are mainly concerned with implications and effects on InfoVis and VA procedures, we examine properties of seven dimensions that constitute the main characteristics of a visualization pipeline. As illustrated in Figure 3.3, we adopted the VA pipeline by van Wijk et al. [380]. We consider their model to be particularly suitable for our needs, as they provide a highly generalizable pipeline with the focus on the "user side" (knowledge generation, interaction, perception, etc.) instead of the "computer side" (parameterization, model evaluation, etc.). Therefore, we use the states and transitions they identified in the pipeline to compare two different media that display the same visualization. We compare the most established medium used for InfoVis and VA tasks – the monitor screen – with the ideal "VR-display" described in the previous section.

(i) Data Input of the information visualization pipeline (Figure 3.3 (i)). Assuming that the data collection process is not a visualization task itself, this dimension is not directly influenced by the medium used. However, some data types are predestinated for specific visualizations, and their performance could be affected by the medium used.

(ii) Visualization Data transformation and visual mapping steps, as they are described by Card et al. [395] in their well-established reference model for visualization, are covered by this dimension (Figure 3.3 (ii)). It describes how the specification is applied to the data (e.g., how the data is transformed, how visual variables are mapped). The outcome is an “Image” that can be perceived by the user. The visualization is displayed on the virtual monitor in the VRE – just exactly the same way as it is shown on the real monitor. Any data transformation, visual mapping, etc. applicable to the “real” visualization on the monitor screen would also be applicable to the VR visualization.

(iii) “Image” (Image, Animation, etc.) Outcome of the visualization step (Figure 3.3 (iii)). This dimension describes properties of the actually perceptible entity that is observed by the user. In the field of information visualization, this is most often a 2D image, but it could also be audio, video or haptic feedback, or any other perceivable signal. The process of transforming the visualization into a user-perceivable signal (“Image”) constitutes the view transformation step in the model by Card et al. [395]. Applicable types of “Images” depend on the specification given in a respective analysis task. Vice versa, hardware specifications, such as the medium used for display, define characteristics of the “Image”. Assuming that the VRE is ideal, the image presented to the user is the same as in the real world.

(iv) Perception & Cognition Conceiving the “Image” (Figure 3.3 (iv)). This dimension describes how the visualization is perceived and processed by the user. This step is user-, visualization- and specification-dependent. Different users may perceive the “Image” differently, for instance, due to diverse personal notions (e.g., interpretation of “Image”) or physical differences (e.g., color perception). Obviously, also the type of “Image” (e.g., line-chart visualization) and its specification (boundary conditions such as the hardware used for display), influence its perception and cognition. In an ideal VRE, limitations like low resolution, latency, or the weight of a head-mounted display are neglected. In this case, the VRE and the real environment cannot be visually distinguished from each other. Therefore, perception and cognition in the VRE would on all channels be equal to a real-world experience.

(v) Knowledge Outcome of the visualization pipeline: the knowledge gained in the course of the analysis (Figure 3.3 (v)). Its volume and character depend, among other things, on prior knowledge (e.g., experience, domain-specific knowledge) and the expressiveness of the inspected visualization. The latter depends on several properties, such as the characteristics of the underlying data, the type of visualization used, and the aspect of interest. In an ideal VR, everything is cloned perfectly, leading to the same situation in the virtual and real environment. Therefore, perceiving a visualization in the ideal VRE instead of inspecting it on a computer screen would not affect the knowledge generation process in any way.

(vi) Interactive Exploration & Collaboration Types of interaction and their properties (Figure 3.3 (vi)). This dimension describes where and how users can interact with the system (i.e., adaption of specification) or with other users (collaboration). In our thought experiment, the real world is cloned, and everything in the virtual copy can be touched and interacted with. Interaction would be the same as in the real world. Collaboration could be achieved, for example, by screen sharing (of the virtual monitor screen).

(vii) Specification Boundary conditions for the creation of the visualization (Figure 3.3 (vii)). The specification defines properties like the type of visualization, the dataset to be considered or the mapping of data attributes to visual variables. Moreover, this dimension describes the hardware and (physical/virtual) environmental conditions. Data processing is covered by this dimension as well: Preprocessing, such as normalization or data cleansing, as well as actual data mining steps, such as clustering, pattern detection or classification are implicitly reflected in the pipeline. Such operations can be seen as the execution of an adaption of the specification that leads to a manipulated dataset. For instance, in ML/VA workflows, the specification includes any parameterization for the training of the ML model (learning rates, activation functions, etc.). The specification can be adapted by the user through interaction. Most aspects of this dimension are independent of the medium used for the final visualization. Even though the hardware on which the visualization is displayed in the VRE is different, the outcome is the same.

3.2.3.2 Renouncing the Ideal

In this section, we will qualify statements made in section 3.2.3.1 with respect to actual limitations of VREs. Due to technological flaws, the state that a VRE is perceived in exactly the same way as the real world may never be reached. However, for visual analysis, a VRE can already be considered “ideal” if it lacks nothing that would compromise the efficiency or effectiveness of the visualization or task. Therefore, the simulation of the real world does not necessarily have to reach a perfect level at which the VRE cannot be distinguished from the real environment.

Starting with the specification, it would never be possible to perfectly replicate a real-world scenario in VR without any flaws. Detail, resolution, photo-realistic rendering, real-time monitoring, and cloning of real objects into the VRE are only some examples of the many obstacles. More realistic would be an environment that is anchored at some points to the real environment, but is largely independent. The previous example in the thought experiment could, for example, be reduced to a VRE that only overlaps with the real world in certain parts – the desk and the input devices (mouse and keyboard). Anything else in the VRE could be artificial without directly affecting a particular visual analysis task. Although we would not be able to track remaining physical objects or receive haptic feedback from virtual objects in the VRE, objects would still respond to the input devices used – similar to desktop systems.

The ‘workaround’ used in the thought experiment to deploy hardware input devices for interaction is not the most obvious. Of course, this may be possible, as several prototypes on the market show (e.g., Logitech [228]), but there may be different interaction possibilities that work better with VR (e.g., the mouse only allows navigation on a plane). Mouse and keyboard are the two most established interaction devices most people are familiar with. Introducing new, VR-optimized devices, such as the Vive controller [160], could result in comparatively high learning effort. However, both types of input devices are possible for use in VA workflows. The main goal is to optimize the visual analysis by finding the most suitable interaction device for a particular task.

Depending on the specification, the generated “Image” would be subject to low resolution, non-photo-realistic rendering, and many other artifacts that affect later perception and cognition. The extent to which these artifacts influence the expressiveness and effectiveness of the presented visualization largely depends on the specification and the visualization, but also the task and the user. Some technological drawbacks can certainly be resolved by future technology (e.g., low resolution, latency, photo-realistic rendering). However, it is highly unlikely that one day, the human brain can be completely deceived. Nevertheless, we argue that perfect delusion is not necessary to achieve an equivalent performance in VR compared to screen-based media. The information displayed only needs to be conceived similarly efficient, complete, and exclusive.

3.2.4 Breaking the Curse with Virtual Reality

After having established in the previous section that VREs can theoretically perform just as well as conventional media in terms of conveying information, we attend to the advantages VREs can have. Figure 3.4 depicts two sides of VREs. The properties discussed in section 3.2.3, which compensate conventional media, are shown in red. This section focuses on the blue side, presenting exclusive features of VREs that can have a positive impact on different dimensions. Using the previous example, one could clone the real environment and interact with it as if one perceived the real world. This would not make any difference to interacting in and with the real world directly, *but* it would also be possible to do things which are not possible in the real world. For instance, one could manipulate objects in the environment, install "holograms" or 3D visualizations within the room or exclude distracting parts. In the following subsection, we deal with each of the previously determined dimensions (i)-(vii). Subsequently, we discuss various dimensions that are exclusively available in VREs with regard to their ability to aid in breaking the curse of visual data exploration. This means properties of VREs that help to identify data dimensions that were neglected in the visual data analysis process but have implications for the knowledge generation process.

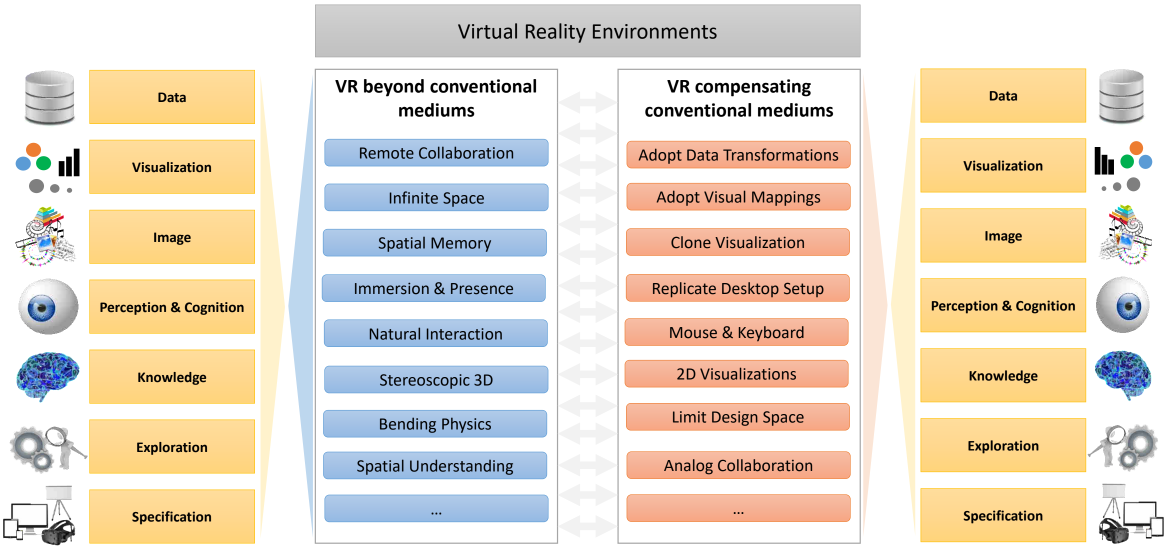

Figure 3.4: We investigate if all properties in the visualization pipeline (yellow) for the display of visualizations on a conventional medium (e.g., monitor screen) can be compensated by VREs (right side, red). This is exhaustively discussed in section 3.2.3. Red lines indicate to which dimension(s) the examples can be assigned. Subsequently, characteristics of VR that are not present in conventional media are discussed in section 3.2.4 (left side, blue). Blue lines indicate which dimensions are affected by the exemplary VR characteristics listed.

3.2.4.1 Implications for Dimensions

Data (i). The "Data"-dimension is not directly influenced by the medium used for display. Nevertheless, some data (e.g., spatial, 3D, volume) can be displayed more naturally in VR than in other media. This means that different data types allow different visualization applications and influence their effectiveness and expressiveness. As shown in the thought experiment, VR cannot only implement any conventional visualization mapping but also extend it by using, for instance, stereoscopic 3D vision.

Visualization (ii). Previous research has shown that it can be advantageous to use the immersive characteristics of VR to combine multiple views that exploit the three-dimensional aspects of such an environment [382]. The user can work with visualizations in the traditional way, while it is also possible to seamlessly combine both concepts in a single application, lowering the bar for users who are not used to VREs. Thus, the analyst is provided with a flexible system that enables him or her to investigate information as intended. For example, Huang et al. [162] combined GIS and VR in an internet application to support users in decision-making processes by deploying spatial databases and spatial data visualizations. Besides the option to map traditional visualizations of a two-dimensional setting into a VRE, it is possible to enrich the presented information with spatial views. In this way, the three-dimensionality of the visualization design space of the VRE can be utilized to integrate inherent topological information with associated spatial information. Based on these findings, VREs enable the usage of more visual mappings and a larger design space compared to monitor-based 2D applications.

"Image" (iii). If we assume that "Image" is the information of the presented visualization that is sensorily perceived by the user, then the number of possible sensory channels increases in a mediated environment. In addition to the two-dimensional visualization and potential auditory information, we can add haptic feedback, expand it with stereoscopic images, and even include the sense of smell and taste. For instance, haptic feedback can be implemented through a glove that provides advantageous touch feedback [375]. The perceived "Image" is also improved by the stereoscopic nature of identifying the exact position and depth of an object in a VR scene. The integration of three-dimensional auditory information into a VRE allows guiding through 3D visualizations and improves the intensity of the "Image" [235]. Depending on the application, using audio can be more intuitive than encoding all information visually. Porter et al. have shown how to encode spatial direction information in the sense of smell [291]. A VRE that effectively implements such techniques could subconsciously guide users in a certain direction to find the intended results or support the process of exploration. Theoretically, applying methods that stimulate the sense of taste would widen the "Image" space and substantially enrich the VR experience.

Perception & Cognition (iv). A major difference to monitor-based visualization approaches is the complete controllability of the environment, which is presented to a user in a VRE. Studies, mainly from the scientific visualization domain, such as the ones conducted by Laha et al. [215] and Forsberg et al. [293], indicate positive effects of immersive environments taking advantage of this aspect on the analysis efficiency of users in certain tasks. As a side effect, environmental factors of classic desktop environments such as changing lighting conditions, noise interference, or third-party interactions or distractions can be mitigated or even completely prevented. As a consequence, when using VR, a user may be able to focus more on the analytical tasks at hand.

Knowledge (v). Knowledge application and generation are indirectly influenced by the medium used. For instance, when inspecting a geo-spatial visualization on a conventional screen, the user would perceive a 2D projection of the scene, whereas he would perceive stereoscopic 3D images when inspecting the same scene in VR. Perceiving the least abstracted copy of the familiar real world (stereoscopic landscape) may help to quickly transfer and apply prior knowledge to the visualization. Vice versa, the less mental mapping is required from the visualization to the real world, the easier it is to transfer gained knowledge to the real world. The smaller distance between the visualization and the real world could ease the transfer of knowledge to the visualization, increase the understandability of visualization results, and thus augment the efficiency and effectiveness of visual analysis procedures. Previous studies (e.g., [210, 245, 300]) have shown that increased immersion can have positive effects on memorability as well. Spatial memory is improved by heightened spatial understanding and muscular memory. In visual analytics tasks, this could enhance efficiency (e.g., cluster identification, analytic provenance) and effectiveness (volume of knowledge extracted).

Interactive Exploration & Collaboration (vi). Traditionally, one can distinguish between remote collaboration and co-located collaboration. Using VR technology, it is possible to simulate co-located collaboration even though the collaborators are physically separated. This can be realized by projecting the avatars of all collaborators into the same VRE. This approach combines the advantages of both kinds of collaboration: The benefit of remote collaboration of not having to be present in the same room, and the improved interaction possibilities of co-located collaborations. For example, one can point at a specific position to guide other persons' attention. On a monitor, the same can be achieved by the use of a mouse pointer. However, this is limited to pointing at objects shown on the screen itself. Pointing at dashboards, persons or hardware other than the monitor is not possible with the mouse. Another advantage of remote collaboration in VR is that it is possible to see where the other person is looking at. Seeing what collaborators are currently focusing on can help to improve the common conversational grounding.

Specification (vii). In comparison to traditional media, one of the main advantages of VR hardware is the fully controllable stereoscopic environment. Visualizations which use a 3D structure to encode information can make good use of this as their shape is easier to identify in a stereoscopic view than in a screen-based 3D view. Compared to other media that offer a stereoscopic view, like a 3D print or 3D glasses, VR allows to control and influence the entire visualization space in various ways. This increases the options for displaying information and reduces distractions by removing unnecessary elements from the scene. This works well together with the large field of view that VR hardware usually offers. A large field of view allows showing more information at once since more graphical elements can be seen at the same time. Finally, head-mounted displays are easy to transport, which can be an asset in some cases. Other media that offer a large field of view, for example, large-scale displays, are often by far more difficult to transport or not transportable at all.

3.2.4.2 Minimizing the Gap between the Real World and Data World Using VR

As shown in the introduction of the curse of visual data exploration, it would be optimal for the validity and applicability of analysis results if the discrepancy between the data world and the real world would not exist. In reality, however, this is hardly possible, since the digitization process of real-world properties is already subject to a loss of information. Nevertheless and as previously shown, the use of VR may minimize the gap between the data world and the real world as it allows to inspect the data world more naturally and realistically than on a conventional screen. In VR, the snippet of the real world can be inspected on a lower level of abstraction due to stereoscopic 3D perception, natural navigation, and improved spatial understanding. Because of immersion, differences between the inspected virtual environment and the familiar real world can be identified more easily. For instance, one is automatically aware of the absence of familiar properties, such as the atmosphere, wind, sound, and smell.

For the validation of analysis results, the real world could be reconstructed from the data world as close as possible. Due to an incomplete data world, the reconstruction would suffer from incompleteness as well, calling attention to missing or faulty dimensions in the data. Subsequently, analysis outcomes could be projected into the reconstruction. The analyst could then enter the resemblance of the real world and verify if displayed analysis outcomes logically fit in the displayed environment. For instance, based on experiences from the familiar real world, the user would be able to identify errors in analysis results, such as an outlier in a person's movement trajectory that describes a jump over five meters and back in two seconds. Moreover, the absence of important dimensions that possibly influence the analysis outcome could be identified and included in the next analysis iteration. For instance, in a collective behavior analysis task, tracked animals could be represented in a virtual reconstruction of the landscape in which they actually moved. By inspecting this virtual scenario, the analyst may be able to detect errors in the analysis. For example, analysts may recognize that they have not considered obstacles such as rocks, trees, or rivers that influence the animals' trajectories. Animal paths may have certain shapes due to characteristics of the environment, which are easily overlooked in a basic analysis without a more detailed visual examination. VR could also help to identify dimensions that are more difficult to detect on a screen. Being immersed in a copy of the real world allows the analyst to perceive the environment more naturally. Artifacts that contradict the familiar real-world properties, such as gravity, atmosphere, or three-dimensionality, may be identified more quickly than without immersion. For example, when analyzing collective behavior of baboons [361], often only the 2D trajectories of the animals are considered. When representing the trajectories in most detail (3D) in a virtual environment, the analyst would automatically become aware of the circumstance that the animals tend to climb tall trees overnight.

Previous studies have shown a benefit of immersion and VREs for spatial understanding [12], depth perception [6, 387], and spatial memory [67]. For example, in soccer analysis, when standing immersed on the virtual soccer field, distances can be estimated more accurately due to stereoscopic perception and familiar distance estimation, since everything is presented according to reality. VR is accompanied by further benefits, as discussed in the previous section. For example, it provides a platform for natural remote collaboration, the entire space in the environment can be used for the display of visualizations, and laws of physics can be bend.

Disclaimer. We do not argue that VR is the one and only solution to minimize the gap between the data world and the real world. For many abstract data types, there is no straightforward mapping to the real world, which allows a visual validation of the real-world context. For instance, when analyzing stock exchange rates, there is no reasonable way to connect the numerical time series data to a physical real-world location. It would not make sense to map analysis results, such as parallel coordinate plots with highlighted clusters, into the virtual duplicate of a trading floor. To apply our approach, a direct connection between the data and the real world is necessary. Moreover, there are many visual analytics applications in which no sufficient data basis exists to reconstruct the real world from the data world and adequately display the analysis results in it. However, there is a wide range of domains in which VR could be used to minimize the discrepancy between the data world and the real world, such as in the emerging fields of collective behavior analysis [43], sports analysis [356], and general geo-spatial data analysis [252]. In the following section, we present an example where we deploy our solution approach for a criminal investigation use case.

3.2.5 Use Case: Crime Scene Investigation

In this section, we discuss how virtual reality is used in a visual exploration workflow deploying a use case from the field of crime data analysis. To detect differences between the data world and the real world (the curse of visual data exploration), we try to reproduce as much as possible from the real-world scenario in a virtual environment. Analysis results are, additionally, embedded in the reconstruction to be verified and analyzed by a domain expert. We first describe the developed prototype. Subsequently, we focus on the two integrated solution approaches to "break the curse" by reconstructing the real world from the data world and projecting the analysis results into the real world. Afterward, we discuss the role of VR and its benefits in this use case. Last but not least, we provide results from initial expert feedback on the expansion of the screen-based system with VR.

3.2.5.1 Prototype for 4D Crime Scene Investigation

After a criminal act, police agencies collect all obtainable information to reconstruct the course of events in the minutest detail. Witness videos and surveillance footage are often central sources. The project FLORIDA [118], as well as the EU project VICTORIA [383], focus on facilitating the analysis of large amounts of video data for LEAs. The FLORIDA project is part of a bilateral project of the German BMBF and the Austrian Security Research Programme KIRAS, which is funded by the Austrian Ministry of Transport, Innovation and Technology (BMVIT), and is run by the FFG (Österreichische Forschungsförderungs Gesellschaft). In the scope of these projects, a tool is currently being developed that allows analysts to merge multiple video sources into a single timeline. Therefore, we visualize a 4D scene (3D space + time) from all input videos by aligning the video sources spatially and temporally. In addition, we enrich the scene with analysis results of semi-automatic object detections from machine learning algorithms. That way, large sets of videos can be inspected simultaneously. Figure 3.5 shows a sample crime scene (left), an exemplary frame from one of the installed surveillance cameras (center) and a point cloud reconstruction of the crime scene (right). The described use case is based on the multi-camera dataset "IOSB-4D" provided by Pollok [288]. The reconstruction and automatic detections were provided by project partners. Our contribution to the framework is the visualization of all provided information in a (VR) VA environment. In the following, we will explain the framework and its usage.

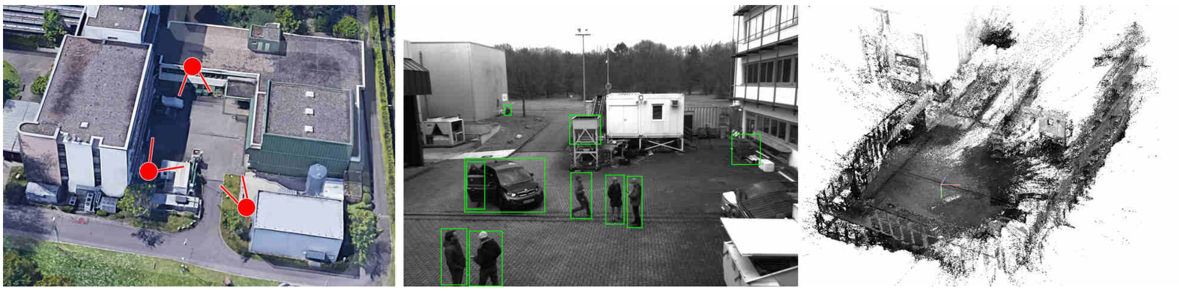

Figure 3.5: Crime scene reconstruction. A crime scene (left) is monitored by multiple video cameras (red dots). Each camera provides video footage (center) that can be analyzed with machine learning algorithms (bounding boxes indicate object detections). Using multiple video sources, a 3D point cloud can be reconstructed from the 2D video streams (right).

Input Data. The data base for the described analysis is a series of surveillance and witness videos that recorded the same incident from different angles. The sources can be statically installed surveillance cameras or moving hand-held devices, such as cameras or mobile phones. Figure 3.6 depicts the merging of all video sources into a dynamic 3D scene that is enriched with automatic detection and analysis results, which are also based on the input videos.

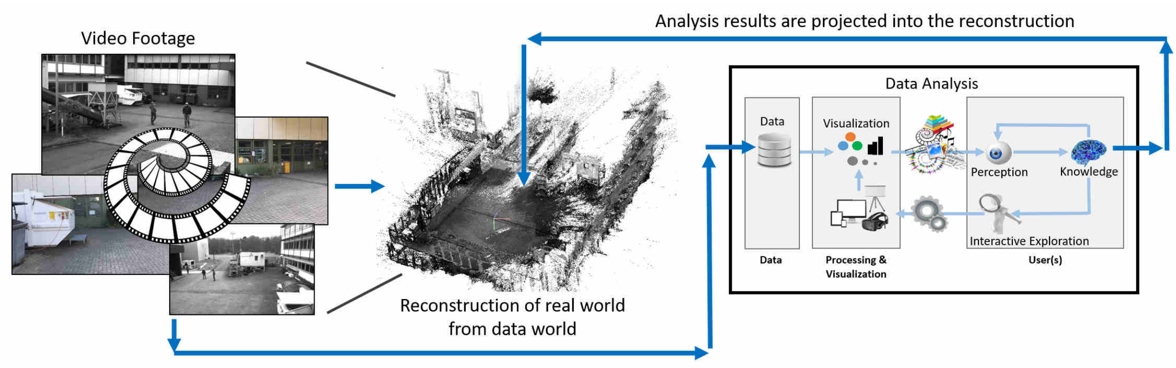

Figure 3.6: Multiple video sources are positioned in a 3D scene. With feature matching algorithms, a 3D scene is generated from 2D videos. The origin of each video is registered in this 3D scene. In addition, the videos are analyzed individually, and the results from semi-automatic object detections can be visualized in the 3D scene.

4D Reconstruction. First, the videos are synchronized temporally by aligning them on a common time axis based on timestamps in the videos and striking events within the videos. In a second step, all video sources are spatially aligned using feature matching algorithms. As a result, for each timestamp on the common time axis, the origins of all video sources are positioned in a 3D scene. Next, visual features in individual video frames are manually geo-referenced to transform the 3D scene in metric space. For stereoscopic cameras and depth cameras, the videos are transformed into dynamic point clouds as depicted in Figure 3.7. Additionally, a static background mesh can be created using a structure from motion approach (see Figure 3.8). The credits for the reconstruction go to our project partners from Fraunhofer IOSB.

Figure 3.7: Dynamic point cloud visualization. Each frame in the source video is displayed as a 3D point cloud. The 3D video stream can be played or skipped through by using a timeline slider in the interface.

Video Analysis. The input videos are processed in a semi-automatic feature detection procedure. For each video, objects and persons are detected and tracked for the entire duration of the video. Hereby, pathways of detections are created that can be positioned and visualized in the 3D scene using triangulation (see Figure 3.8).

Usage of Tool. The main purpose of the developed tool is to support LEA officers in their investigations of crime scenes. The key benefit is the capability of the tool to visualize multiple source videos at once and in the correct spatial and temporal context. The user can explore the large set of witness videos by browsing through the 4D scene. But the tool is not only suited for the inspection of raw footage videos that are arranged in a 3D space. It can also be used for more advanced visual analytics procedures. For instance, results of object detection algorithms can be classified in different categories (e.g., person or car) and displayed as heatmaps (Figure 3.8, bottom line). This allows the user to quickly grasp the overall distribution of the occurrence of certain objects in the entire scene. The user can refine the visualization by filtering the displayed domain or selecting some classes to display them as multiples of heatmaps. Another example is the analysis of trajectories. Movements of persons and vehicles can be tracked and analyzed. The analysis progress follows the visual analytics pipeline model depicted in Figure 3.3. The analyst can select multiple time series to compare, apply operations to them (e.g., clustering, classification, event detection), display intermediate results, refine input parameters, and continue this loop until the results are satisfying. Subsequently, the outcome can be presented in the 4D scene for verification and inspection.

Figure 3.8 shows the interface of the 4D crime scene investigation tool. A timeline at the bottom allows for the temporal navigation within the incident. For spatial navigation, the analyst can fly through the virtual space using a keyboard. A mini-map on the top-right provides an overview of the scene and depicts all registered surveillance cameras in yellow. The original video footage can be inspected by clicking on one of the registered cameras. Detected objects are displayed in the 3D scene as snippets from the original video material, framed by green bounding boxes.

For the evaluation of evidence footage, it would be possible to dispense from visual analysis and only evaluate the results of object detection algorithms quantitatively. However, visual exploration of video sources combined with automatic detections allows domain experts to identify false classifications, outliers, unrecognized entities, and relationships between subjects that are not visible in the mere metadata from automatic procedures. The merging of all video sources into one 4D scene allows the analyst to quickly capture the entire data space without having to watch each video separately. Moreover, the spatial and temporal alignment of video sources helps to understand the spatial and temporal connection between different video sources. The presented approach combines the two presented strategies to bridge the gap between the data world and the real world [206]. The real world is reconstructed from the data world by creating a dynamic 3D scene from 2D video footage. Additionally, analysis results (machine learning object detections) are visually projected back into this reconstruction of the real world (see Figure 3.5 center).

3.2.5.2 Reconstruction of the Real World

There is always a discrepancy between the real world and the data world on which the analysis is based on (see section 3.1). This gap may lead to wrong assumptions about the real-world scenario. In the present case, this may be due to dimensions not considered for analysis or incomplete video sources. For example, the scene may contain spatial or temporal black spots that were not monitored but are highly relevant for the reconstruction of the entire sequence of events.

The reconstruction of the real-world scenario from the data world provides an overview and eases the identification of errors in the data. It may also reveal dimensions that have been neglected in the analysis. If inconsistencies between the data world and the real world are found, they can be resolved,

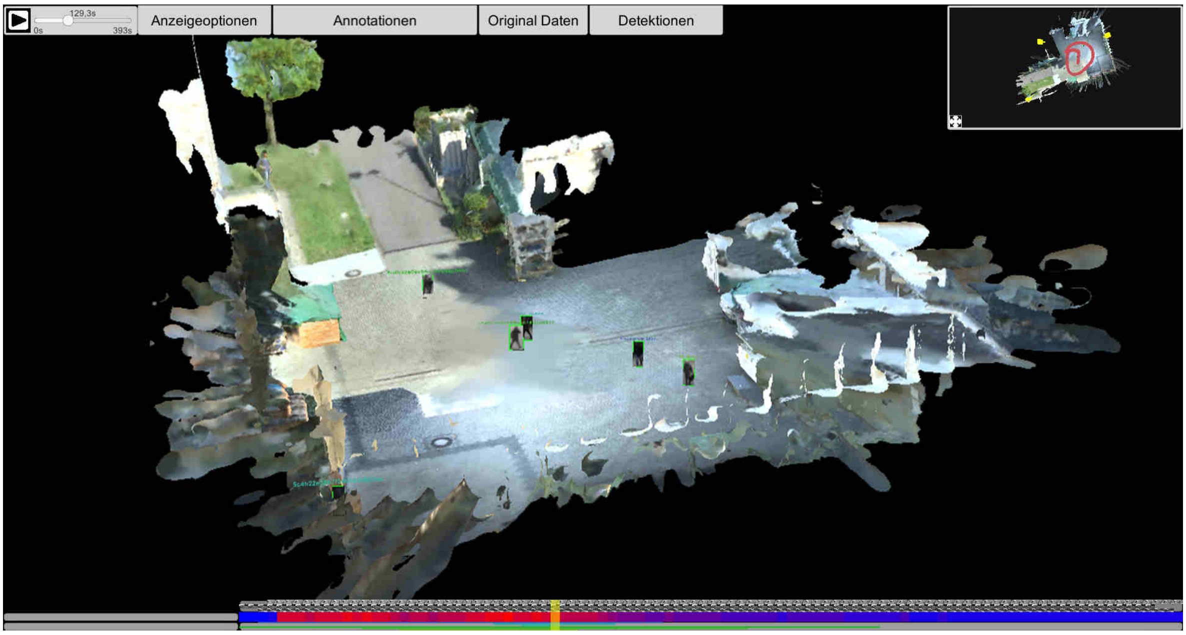

Figure 3.8: Reconstruction of the real world from the data world and projection of analysis results. The generated 3D point cloud is converted to a 3D mesh of the crime scene. Object detections from all video sources can be displayed within the scene (green rectangles with snapshots from videos).

which leads to a minimization of the gap. By merging all available video sources into one 4D scene, a common frame of reference is established that maximizes the context of each individual video source. For example, when inspecting one video in the 4D scene, the analyst is aware of all sources that monitor the same area of interest and can switch between multiple videos without having to search the entire pool of footage. This connects all available sources and facilitates the overall analysis. When inspecting the virtual reconstruction of the environment, the analyst subconsciously relates perceived information to known real-world properties. Hence, not modeled dimensions can be identified that have also been neglected in the analysis of individual video sources. For example, a scene may appear unnatural if there is no sky with sun, indicating that weather conditions were not taken into account in the analysis.

3.2.5.3 Projection of Analysis Results

The basic idea of this approach is to verify the data of the analysis as well as the analysis outcomes by projecting the results of the analysis into original real-world footage. Displaying analysis results in original footage is less abstract than simply presenting the results without context. For example, when analyzing trajectories of soccer players in a soccer match, it may be more useful to visualize the extracted trajectories directly in the video from which they were extracted than to create an abstract replica of the soccer field for visualization [206]. In the current use case, the 3D reconstruction of the scene can be used to display spatially and temporally aligned automatic detections. This alignment helps to verify intermediate analysis results, such as automated machine learning outputs. Hereby, falsely classified objects can be easily identified.

Moreover, the projection of analysis results helps to associate detections of the same entity coming from different video sources. For example, if the same person is detected in multiple videos but not matched, this results in multiple individual detections for the same entity. Displaying extracted trajectories that describe the movement of entities can be visually verified. In the automatic extraction process, it often happens that affiliations of trails are mixed up when the trails of two entities cross each other. For example, if two persons go to the same location and move apart again, the trail of the first person is connected to the trail of the second person at the point of encounter and vice versa. Visually, this flaw can be identified easily, but it is hard to detect it in more abstract representations of the analysis results. Additionally, it can happen that the pathway of a single entity is split into multiple segments and is, therefore, considered to be from different entities. By visualizing the segments in the virtual environment, such segmentation can be identified and resolved by merging them into one pathway.

3.2.5.4 Advantages of Deploying Virtual Reality

Virtual reality allows the analyst to dive into the crime scene and to be fully immersed in the progress of events. The scene can be inspected "from the inside" as if the analyst had been present at the time of the incident. The course of events can be followed in a 3D video of the incident, reconstructed from all available footage. The real world is better replicated since the analyst can naturally walk through the scene. Details, such as the absence of wind or sound, may come to the analyst's mind as the reconstruction is still dissimilar to the known real world.

The reconstruction of the real world from the data world is implicitly improved by changing the way it is observed. The validation of analysis results can be enhanced by augmented depth perception and stereoscopic 3D vision. Distances may be perceived more accurately, allowing a more precise assessment of displayed trajectories. For example, if two waypoints of the trajectory of a person in the scene are far apart, being immersed can help the analyst to classify the respective segment of the trajectory as erroneous. The analyst perceives the virtual environment in familiar metric 3D space (distances are measurable in meters), allowing the user to judge properties based on personal experiences from the real world. For instance, it may become obvious that it is not possible for a human being to cover the depicted distance in the given time.

Various studies have shown a positive effect of immersion on spatial memory [27, 50, 100]. In our use case, too, the criminal investigator may benefit from immersion with regard to navigation and orientation capabilities, and it may help the analyst to keep track of suspicious actualities in the scene. This is also fostered by how the analyst can explore the scenario. Instead of having to use tools like mouse and keyboard, the analyst can walk naturally through the virtual environment and focus on the analysis task, not on navigation and possibly complex interaction.

The deployment of VR poses several additional opportunities. Among others, it enables remote collaboration of multiple LEA officers in the shared crime scene reconstruction. This way, a common conversational grounding is established, and the communication is eased by displaying avatars for each participating analyst [21]. The avatars can be used to perform gestures and to call attention to something by pointing at it. Moreover, the VRE offers plenty of space for the installation of infographics, control panels, and interfaces for the overall analysis.

3.2.5.5 Initial Expert Feedback

The presented tool was developed for LEAs to support them in the investigation of crime scenes. To receive initial expert feedback, we surveyed eleven criminal investigators (CI1-CI11) from the German federal police (Bundespolizei). All invited criminal investigators are employed as police inspectors (degree A10-A15) and have between 10 and 35 years of professional experience. All investigators work with video surveillance or video evidence evaluation on a daily basis. First, we demonstrated

the capabilities of the tool and gave a tutorial on how to use it. Then, each expert had the opportunity to use both the standard tool on the screen and the VR extension. After that, we conducted a short interview with each participant to obtain qualitative feedback on our approach. The interview consisted of three parts. First, we addressed the crime scene reconstruction. Secondly, we put the focus on the tool's capability to display analysis results, such as object and person detections. Thirdly, we discussed the use of VR and the challenges and benefits it brings.
The results of our initial evaluation are very promising. Investigators acknowledge its potential to be deployed in the fight against crime, among others as a supporting tool for surveillance tasks, recon missions, and digital forensics. VR, in particular, was perceived as having huge potential in this specific domain. VR was even described as a quantum leap forward in criminal investigation by one of the interviewees.

Reconstruction of crime scene. The reconstruction of the crime scene is perceived as very helpful for investigators to re-enact the crime by providing a visual basis for imagination (CI7). Image-based reconstruction from evidence footage is not only faster, cheaper and more flexible than advanced laser scans of the scene in the retrospective forensic procedure, as it is current state-of-the-art in this domain, but also provides a snapshot of the environment from the time of the incident (CI5). Thus, situational conditions which were present at the time of the incident but changed during or shortly after the incident would not be present in a post-event recording of the environment (e.g., when the scene is scanned with lasers after the incident took place). Moreover, the detailed reconstruction of the crime scene puts all video sources into context, promotes orientation, and helps video analysts to put multiple video sources into a temporal and spatial relation (CI5, CI6). Another benefit mentioned was the versatility of the approach. Analysts do not have to physically go to the crime scene but can inspect it from anywhere (CI9). This may also be used in recon missions to gain orientation in an unknown environment before police forces are deployed. As opposed to watching video sequences from the environment, a 3D reconstruction improves spatial understanding (CI5, CI9, CI10). Currently, digital forensics makes use of 360° photo spheres to record characteristics of a crime scene. According to one criminal investigator interviewed, distances and sizes are better perceivable in a 3D scene (P7) than in image spheres or videos. Many of the LEA officers saw the need for a more photo-realistic reconstruction and expect great benefits (CI1-CI4, CI6, CI9, CI10). With advancing technological progress, we suggest that the quality of image-based reconstruction will improve significantly shortly and, therewith, minimize this constraint.

Embedding of analysis results. A major benefit expressed by the interviewees was that insights from all video sources are merged into one common scene. Not only are the cameras referenced locally in the scene, but also are the detections of their footage set into a temporal and spatial relation. This makes it easier for the investigators to get an overview of the scene and to re-enact the course of events (CI1-CI3, CI8). The reconstruction in combination with automatic detections from the video footage sets all analysis results into context and creates a "big picture" of the scene (CI7). The detections themselves lead the investigators to interesting and important events and characteristics in the huge pool of evidence videos (CI4, CI6). Often, videos are manually inspected, frame by frame, for hours in order to detect wanted persons or vehicles. The automated approach draws investigators' attention to areas in a video that contain movement or even certain types of detections (persons, faces/identities). Moreover, it facilitates the combined inspection of multiple video sources. For example, it allows to easily identify which cameras monitored which area or person at which time (CI4, CI6). Furthermore, the system can be used to verify automatically generated information, such as detections of persons and movement trajectories. When visualizing analysis results, the analyst can validate the information displayed by relating it to the respective context (environment and detections from other cameras) and real-world experiences (CI9). Moreover, it would be beneficial to ease the mental mapping from the data world to the real world for the display of automated detections, if they would be visualized more realistically (CI10). For instance, detections of persons walking in the scene should be displayed as 3D models instead of image snippets moving around.

The deployment of VR. The interviewed investigators mentioned a strong and clear benefit of VR with regard to its capability to project the analyst into the scene. According to one criminal investigator, VR allows to “dive deeper into the crime scene” as “one forgets the physical environment” (CI8). Thus, the 3D scene in VR is more comprehensible, and spatial relations are easier to understand than when inspecting the scene on a screen (CI1-CI3, CI10). Distances can be interpreted more naturally and intuitively (CI1-CI3). As further benefits and future areas of application, the experts mentioned its potential for remote collaboration and mission training in unknown areas (CI1, CI3). As for drawbacks, the investigators mentioned a high learning curve for the new medium and difficulties with the new interaction techniques with controllers (CI9, CI10). Moreover, they saw possible limitations in the high acquisition costs for the required equipment. We argue that this constraint is just a matter of time as it can be expected that prices are going down since VR devices become more and more established.

3.2.6 Discussion

We have analyzed the applicability of VR for breaking the curse of visual data exploration on a theoretical level. We first showed the theoretical equality of VR to conventional media by going step by step through a VA pipeline model. Thereby, we discussed on what properties a VRE has to fulfill to keep up with a monitor screen and identified characteristics of VR that can pose a benefit. Subsequently, we put the focus on VR properties that can help to break the curse of visual data exploration and demonstrated its feasibility in a use case. However, the approach from the use case cannot be applied to any analysis case. For instance, in order to reconstruct the real world from the data world, there must be an analysis scenario that has an unused visualizable context (e.g., a geo-spatial context) and enough data to actually perform a reconstruction. The strategy described is quite costly, and the potential benefit should be estimated in advance. In some cases, it may be sufficient to rely on alternative strategies to verify the accordance of the data world and the real world for the respective analysis. For example, projecting analysis results back into less abstract data can already help to detect misclassifications or discrepancies in the data (e.g., automatic object detections in videos are overplotted in the original videos). For more abstract data, such as stock market developments, the curse of visual data exploration must still be considered, but assessed formally, for instance, with a detailed influencing factor analysis. VR should only be used if the thereby introduced medium has a significant benefit compared to another medium. For instance, in geo-spatial visualizations, it may be useful to deploy VR due to its capability to project the analyst into the scene, which increases immersion and allows the analyst to inspect the scene naturally.

Currently, state-of-the-art VR technology is placed at a disadvantage because of several drawbacks. Hardware constraints – for instance low resolution or high latency – can introduce physical or psychological phenomena, such as motion sickness or excessive cognitive load. Users are more familiar with non-VR environments, which could affect their overall VR performance. Moreover, many tasks, visualizations, interaction methods, etc. were optimized for conventional desktop PC setups while interaction methods for VR have not yet been sufficiently studied. However, we live in a dynamic era of rapidly evolving technologies, and technological progress may eradicate many of the factors mentioned above shortly, e.g., by providing higher resolution, photo-realistic rendering or advanced interaction technology.

When it comes to the choice of the medium used to display information, it is all about optimizing the visual analytics procedure. According to Chen and Golan [61], this can be seen as an optimization of the cost-benefit ratio. Chen et al. [60] specifically target the cost-benefit analysis for visualizations in VREs. As pointed out in previous work [206], it is important that the analyst can verify the applicability of knowledge generated in the VA process to the investigated real-world scenario. We have demonstrated in the previous sections, how VR can aid such a verification process. However, it depends on the given analysis scenario how the cost-benefit ratio changes when switching from a display-based medium to VR. We argue that the decision to use a VRE as a design space for visualizations must be justified by improvements induced by VR. If the cost-benefit ratio is not affected, we consider it advisable to stick to other media.

3.2.7 Conclusion

We identified seven dimensions which can be used to compare visualizations presented on different media. More precisely, we compared the media by means of individual characteristics that are substantial in the visual analytics pipeline. When comparing a generic VR medium with conventional media on a conceptional level, a theoretical equivalence of the media becomes apparent. Of course, this state will realistically never be fully reached. However, VREs have much potential and can be beneficial in several domains. When choosing the most suitable medium for a specific visualization task, the cost-benefit ratio is a key factor which needs to be considered. We expect technological progress to eliminate some of the current disadvantages of VR in future, which may minimize the “cost-side” and in turn strengthen the “benefit-side” of the cost-benefit equation, providing the basis for an extensive application of VR as a viable medium for presenting and exploring data. We presented a use case showing the potential benefit of VR to break the curse of visual data exploration. In a suitable analysis scenario, real-world circumstances to be analyzed can be reconstructed and enriched with intermediate analysis results. By perceiving the reflection of the real world naturally and with all available contextual information, the analyst may be able to identify discrepancies between reconstruction and reality, such as neglected dimensions that are relevant to the knowledge generation process.

3.3 Visualization in the VR-Canvas: How much Reality is Good for Immersive Analytics in Virtual Reality?

As previously discussed, the decision to deploy VR at all should be well considered, but can pose various benefits and opportunities in several areas of application. When deciding to deploy VR for a certain visual analysis application, the realization is linked to many difficult design decisions. For instance, building user interfaces in VR provides many incentives for the desire to replicate the real world within the "VR-Canvas": e.g., three-dimensional spaces and objects, movement, direct interaction, and realistic lighting. While many of these design decisions might, in fact, support learning and provide a strong sense of familiarity, their benefit for effective analytic tasks remains controversial. Similar to how desktop interfaces adapt and *extend* metaphors from the real world, there is a widespread assumption that virtual reality environments will benefit from *not* replicating every part of the real world and instead focus on transcending reality and improving human experience, perception, and, eventually, cognition. In this section, we collect evidence from studies, opinions, and examples to foster the current discussion on how replicating the real world can improve or impede tasks in immersive analytics. To clarify what we mean by "real world", we look at a range of aspects including *spatiality*, *physics*, *multimodality*, and *visual appearance*.

3.3.1 Introduction

Visualization strives to find methods, visual metaphors, and interactive media to optimize the communication of information, the performance of analytic tasks, problem solving, and decision making [397]. Novel technologies for VR have opened up a new space to support perception and interaction with data visualizations beyond the limits of the desktop [56]. Currently, more visualization interfaces for VR are explored and designed, using technology such as HTC Vive head-mounted displays, mature development environments such as Unity and UnrealEngine, and novel toolkits for immersive visualization design [341]. There are many reasons for considering visualization in VR: stereoscopic perception, direct interaction in 3D space, free body movement [72], a 360-degree immersive view as well as a potentially infinite space to place visual elements. As shown in previous research, these introduced properties can improve visual analytics tasks in certain aspects (e.g., immersion: improvements in task efficiency and effectiveness for spatial well-path planning [141]). Due to the beneficial effects of these characteristics, there is a natural desire to replicate the real world in VR. In most cases, these decisions can be a measure to increase familiarity and to support adaption and learning [167].

However, much has been discussed about replicating the real world in virtual worlds (section 3.3.2). For example, Shneiderman argues "*why not making interfaces better than 3D reality?*" [338], while Elmqvist focuses the discussion on visualization and provides explicit guidelines for the use of 3D visualization on screens and in immersive environments [105]. To effectively use VR in immersive analytics, it is crucial to understand the specific affordances and the potential of VR concerning the level of immersion needed to provide users with successful user experience and an effective analytical workflow [385]. These affordances can include both: to adopt the real world and to transcend it by searching for novel solutions that improve perception, cognition, and the performance of analytic tasks in a VRE. For instance, stereoscopic view could ease the recognition of depth in volume visualizations.

Depending on the analytic tasks to solve, user interfaces may use natural, supernatural, or abstract scene objects, resembling reality as far as technology allows.

In the following we collect evidence from studies, opinions as well as examples to enrich the current discussion on how replicating the real world can improve or impede VA tasks in the context of immersive analytics. To structure our discussion about 'reality in virtual reality', we introduce the concept of the VR-Canvas as a conceptual model to think about designing immersive interfaces in virtual reality (section 3.3.3). We further look at different aspects of the real physical world and their possible mappings to virtual reality (section 3.3.4). The considered dimensions of the real world that can be replicated in a VRE may influence VA-relevant aspects in various ways (e.g., an increase in immersion may lead to higher levels of concentration). We discuss them in the light of possible beneficial effects, such as memorability or presence, that may lead to potential uses and applications for visualization tasks. Based on this analysis of the real world attributes mapping, we close this chapter in section 3.3.5 with a discussion of advantages and drawbacks of real world resemblance in VR.

3.3.2 Current Discussion in Interface Design

Replicating reality in user interface is a strong mechanism and has been practiced since the beginning of interface design; buttons, sliders, the entire desktop metaphor including documents, paper bins and work spaces with the ability for direct manipulation, e.g., via drag-and-drop [340]. On the other side, Stuerzlinger and Wingrave [362] discuss how perfect simulation and realistic environments can lead to undesired, increased user expectations of a system. Moreover, metaphors and realistic replications of the real world may eventually, in the long run, be less efficient than proper techniques such as shortcuts [338].

For virtual reality, some guidelines mention to "establish familiarity" [134]. Others, discussed in various blogs [124, 372], include "Make it beautiful" and state that for an increase in immersion breathtaking scenes are advantageous. In fact, much effort has been put into understanding the psychological, perceptual, and cognitive effects of reality, including concepts such as immersion, social presence [172], and direct interaction [337]. In general, a sense of presence involves a sensation on the user's side of being present (spatial presence), and the interactions with other individuals (social presence) [172]. Immersion, in the context of VR, can also be defined as the sense of being present in a virtual environment, e.g., by removing as many real-world sensations as possible and replacing them with VRE sensations [247]. A study by Seiber and Shafer [337], which involved over 200 students, found that controller naturalness and natural mapping already lead to increases in spatial presence in VR, regardless of the display condition (head-mounted displays, standard monitor). Being used to a specific controller further increases the naturalness of an environment. In VREs, one could use one's finger as a pointer in the three-dimensional space, which would be perceived as more natural and therefore increase the presence of a user.

Niklas Elmqvist discusses guidelines and challenges for the 3D visualization of non-spatial data in a blogpost [105]. One of this guidelines reads *"Don't Replicate the Real World"*. He states that the only advantage of increasing familiarity for a user interface by transferring known elements from the real world to 3D controls is insufficient as *"the whole purpose of a computer is to augment [human] abilities and eliminate [their] limitations, many of them imposed by the physical world"*. For instance, in a task that requires high levels of spatial memorability, the user could be forced to walk around if he wants to change his position. This would probably improve his spatial memory. In another scenario, in which spatial memory is not crucial, a teleport function could be introduced that allows the user to change position without walking and minimizing his physical efforts.

3.3.3 The VR-Canvas

Our definition of the VR-Canvas is loosely inspired by the *AR-Canvas*, a canvas-concept for augmented reality describing data visualization in AR compared to the *traditional canvas* for visualizations, such as 2D screens and paper [13]. The traditional canvas constitutes an empty, two-dimensional, monochrome display space. The type of canvas has a strong influence on the type and design of possible visualizations and their respective task efficiency.

In the same spirit, the VR-Canvas is purely conceptual, meant to support thinking about visualization design and the degree of reality in virtual reality: *What is possible? What is desirable? How can we make the best use of the characteristics of the VR-Canvas for visualization and analytic tasks by replicating reality?*

We describe the characteristics of the VR-Canvas with respect to reality as follows: *(i) Spatiality*: The VR environment defines a three-dimensional space with stereoscopic perception. Using suitable controllers or gesture tracking methods, movements and positions from the real world can be mapped to interactions in the VR-Canvas [72]. An *(ii) immersive 360-degree display* provides display space in each possible viewing direction while users' head-movements define their viewing direction. *(iii) Multimodality* includes additional senses beyond visual perception and proprioception, such as haptics and tangibility, sonic information, speech-input, taste, and smell [239].

Each of these characteristics can be used to create a sense of reality in a VRE. At the same time, they constitute the main motivation for the desire to replicate the real world through three-dimensional spaces and objects, realistic scenes, direct interaction, etc. As it is a thought model, the VR-Canvas is independent of any technical aspects, such as specific hardware (e.g., HTC Vive, Oculus Rift, mobile phone), data structures and algorithms (e.g., scenegraph, rendering method) or implementation details (e.g., programming language, rendering engine).

3.3.4 Aspects of Reality

To better deal with the complex concept of "reality", we look at a set of aspects of the physical reality and how they relate to an adaption into the VR-Canvas for visual analytics tasks. Our list is non-exhaustive at the moment, but can be used in future work as a basis to continue the discussion. This list of aspects of physical reality refer to VR properties which may influence characteristics like immersion and presence, which can have an impact on visual analytics tasks (e.g., [141]).

Three-Dimensionality and Stereoscopy Three-dimensional objects and visualizations have largely been condemned for the use on 2D screens since they cause occlusion and perspective distortion, and are hard to interact with [10, 105]. Most of these problems persist within the VR-Canvas. While some studies have shown increased performance in perception for steroscopic virtual displays [111, 242, 400] and physical visualizations [175], others attribute the effect rather to motion parallax [12, 376]. Depending on the data (e.g., sparse, inherently spatial, such as fluids and anatomy) and task – e.g., general overview, convey a metaphor [10], identify purely 3D structures, such as a correlation in three dimensions, or perform complicated interactions that require a high number of degrees of freedom [12]

– stereoscopy might be of limited use. The VR-Canvas [423] can provide alternatives to proper 2D representations.

Laws of Physics Based on physics, humans can infer about objects in the real world, even if they have never seen the object before: gravity and direction of movement, occlusion and positioning, rigidness and elasticity of objects, etc. However, in the VR-Canvas, we have full control over the design space, which allows us to turn off physics for some or all objects. At first glance, disabling the laws of physics in the VR-Canvas may seem preferable as they often limit possible actions (e.g., floating points in a scatterplot visualization). Though, there are problems originating from removing the laws of physics. For instance, in the real world two objects never occupy the same position at the same time. Ignoring this assumption in the VR-Canvas (as well as on 2D media) can cause a person to miss information, such as overlapping points in scatterplots or lines in PCPs. On the other hand, the ability to ignore the rigidness of objects allows higher precision (in scatterplots) and comparing 3D objects through "superposition". The designer should take these effects into account and can offer additional complementary views. It is also possible to enable physics for some objects, but disable it for others, keeping in mind consistency and ways that prevent possible confusion. Depending on how many other aspects of the real world are included in the VR-Canvas, a user may even expect objects to be influenced by physics. If the VR scene only consists of abstract visualizations, it might be easier to accept that they float in mid-air than if the scene emulates an entire office with many objects known from the real world.

Visual Appearance Visual appearance includes everything that influences the appearance of objects in the VR scene: texture, lighting, shading, reflection, etc. Light sources in the scene that create shading on the 3D objects can help the user to identify 3D shapes [368]. This effect can be increased by placing the light source in specific positions, ideally above a user's height [130]. For some tasks it can be advantageous to include light sources that cannot exist in the real world, for example by using some global illumination that makes the whole scene brighter. Lighting can also be used to emulate the day and night cycle in a VR environment. If the user is working at night, it can be confusing if he needs to switch between a dark real world environment and a bright virtual reality environment since the eyes need to adapt to the level of brightness, as is the case with most navigation systems in cars.

If objects from the real world were used in the VR-Canvas, it would be possible to use exact copies of real-world *textures*, which reduces the time it takes to recognize objects [317]. Textures can be used to convey information [234], yet they can make scenes visually complex. To the best of our knowledge, there are no studies to inform decisions about the extent to which texture and shading support or impede analytic tasks in virtual reality.

Environments and Objects The replication of generally physically present objects (e.g., a desk, a room, a tool) inside the VR-Canvas has proven to be beneficial if the physical properties of the objects are maintained (e.g., location, mobility, shape) [342, 386]. Such replication leads to an increased suspension of disbelief in the virtual environment and was used for tangible user interfaces (TUIs) in VR [72]. Slater [346] identifies place illusion and plausibility illusion as two basic factors leading to realistic behavior of users in virtual reality environments (i.e., how real the place and its reaction to user actions is). The sensation of presence (i.e., feeling of "being there") is directly influenced by the two aspects. This is further supported by a study which concludes that "maximal presence in a mediated experience arises from an optimal combination of form and content" [314], i.e., a system has to reflect intended user purposes and mimic expected reactions to convey a high level of perceived presence. However, there is also research contradicting this suggestion. Mental load could be reduced by minimizing the perceivable environment and helping to focus on the analytic tasks.

Navigation in VR space is a common problem as users often need to cover longer distances than they can or want to walk in the real world. While physical solutions exist [85], entirely virtual solutions include teleportation and fast movement. In case these methods are unavoidable, specific care must be taken to prevent motion sickness. Alternatively, spaces could be scaled down or virtual elements could be brought closer to the user, e.g., through pointing, without changing the user's actual position.

Direct Manipulation The well-established device pair of mouse and keyboard is most commonly used for traditional screen-based user interaction. In contrast, there is no golden standard for virtual reality environments. Various techniques have been developed, using tangible controls, such as the Vive controllers, or implicit controls, such as movement or gestures. Wagner Filho et al. [386] purposefully used in their VA prototype a seated setting in combination with tangible controls to reduce the physical effort of the user for spatial navigation in their visualization. This may lead to a decreased feeling of presence due to a less real-world-like navigation, but increases the overall efficiency of the VA procedure.

Collaboration In collaborative scenarios, it may be useful to realistically replicate physically remote collaborators in the VR-Canvas to facilitate more effective collaborations [14, 243]; users can point at objects and convey information through mimic and gesture [75, 126]. In collaborative scenarios, replicating the real world, i.e., at least the collaborators' position, hands, and facial expressions, is highly desirable. This enables the collaborator to use natural communication, such as gestures or facial expressions. Other than in reality, the collaborator does not have to be present in the same physical space as oneself. The use of embodiment as a tool to solve specific tasks was discussed by Mennecke et al. [243] and deployed in many areas, for example, to understand behavior [187] and to raise awareness of social matters [231].

Sonic Information Sonic stimuli convey information in many cases: telephones, alarm bells, Geiger counters, or heart rate monitors. Some sound cues convey information even though they are not designed for this purpose, such as printers and cars. Using the same sound cues that people are already familiar with can avoid additional time for familiarization and provides an intuitive way to convey information to users. Sound can be especially useful to transmit ambient information within a VR-Canvas that supports body and head movement. This allows to divert the user's full attention towards specific elements at a time and the user effectively neglects the rest of the scene which might be behind him or far away. To decide whether a specific sound cue should be included in a VRE, it should be checked whether the sound carries important information and whether this information cannot be transferred more efficiently using visualizations. However, it should be taken into account that a sound can get the attention of users, regardless of the direction in which its source is located. In contrast, visual cues can only be noticed if they are in the user's field of view and may otherwise be overlooked. A comprehensive discussion on multimodality in immersive analytics is provided elsewhere [239].

Metaphors A metaphor is a figure of speech in which a word or phrase literally denotes one kind of object or idea, but is used in place of another to indicate a similarity or analogy between them [92]. Interface metaphors have long been used to embody functionality and familiarity. Transferred to immersive analytics, visualization-, control-, or interaction-elements can be treated as metaphors to control interaction. For instance, one could draw a visualization on a 3D resemblance of a flip chart on a screen instead of depicting it on a modest plane. The visualization could then be hung on a virtual wall, printed on virtual paper, or be organized in shelves.

Even though the use of interaction metaphors, such as walking, grabbing, and gazing, may reduce the learning curve, it can also increase physical and mental effort, resulting in lower efficiency and fatigue. However, in some cases–especially in the context of VR – it may be reasonable to exploit advantages of such metaphors. For instance, by deploying the metaphor of walking (which could be replaced with teleporting in VR), spatial memory could be fostered. By forcing physical rotation for navigation, the sense of orientation could be improved, for instance, by placing items in cardinal directions around the user (who would learn that specific items are always north of him).

3.3.5 Discussion & Conclusion

Our work showcases some of the advantages and drawbacks of real-world resemblance in VR to foster discussion about its benefits and limitations. We only focused on replication, although other guidelines could have as well been considered in more detail. Even though our investigations were motivated by finding a guideline for replicating the real world in visualization tasks, our approach is also applicable to other domains (e.g., gaming, non-VA tasks). We identified eight main attributes to systematically analyze real-world resemblance. Our main concern is when and to what degree the real world should be replicated in a virtual environment for visual analytics purposes. We arrived at a selection of design considerations and recommendations from state-of-the-art related works. The selected papers address advantages and disadvantages caused by the replication of the real world, either explicitly or implicitly. We also analyzed other sources, such as online sources, that discuss design guidelines for VR, 3D and visualizations in general.

Firstly, we suggest that several rules regarding visualization and interface design for conventional media can be transferred to VR. In particular, we refer to the guideline "Don't replicate the real world" by Niklas Elmquvist [105]. When this rule is associated with user interface elements, it leads to a reduction in mental and physical effort to trigger certain actions. It is closely related to a rule established by Shneiderman: "Enable frequent users to use shortcuts" [338]. Thereby, inconvenient realistic actions are replaced by "unrealistic", effortless supernatural equivalents (e.g., allowing to teleport in VR to certain points through a click on the controller instead of forcing the user to walk through the virtual room).

Secondly, we argue that taking advantage of some derived effects of VR, such as immersion, presence, and spatial memory, could be beneficial for immersive analytic tasks, as shown by several prior works (e.g., [111, 141, 242]). Depending on the task at hand, the designer has to balance which aspects have to be replicated more realistically in order to achieve a specific goal and at what price this is done. It is important that the cost-benefit ratio is optimal – i.e., if a realistic replication of the real world is installed and comes along with some disadvantage (rendering effort, distraction), it has to be outweighed by its introduced benefits (e.g., increased level of immersion leading to better performance in a specific task).

Our discussion exposes that it is not generalizable how much reality should be striven for in VREs. Depending on the purpose and impact of possible real-world resemblance, an optimal trade-off between simplicity and realism has to be chosen individually for each application. Moreover, the discussion reveals many points for future research, e.g., to investigate a better understanding of visualization design spaces and to examine how to construct optimal working environments for visual analytics workflows, how to deploy shortcuts in VR, or if the VR-Canvas is suitable for the display of pure 2D contents (which are typical for state-of-the-art VA procedures).

4

Evaluating Human Factors in Virtual Reality

Contents

Many technical and perceptual facets distinguish immersive environments from conventional screen-based analysis environments. In this chapter, we focus on the second strategy for assessing the applicability of VR for data visualization, namely low-level, single-aspect evaluation. To illustrate the practical realization, the assessments of two different aspects are discussed in more detail.

In the first part, the impact of immersion on a cluster identification task on scatterplots is examined, comparing different visualization design spaces on screen and VR setups. This part is based on our InfoVis paper [208], in which we present two user studies comparing different visualization design spaces in terms of their level of immersion and their influence on given cluster identification tasks. The use of immersive environments is associated with various challenges. One of them is the discrepancy between virtual and physical space. While physical space is usually very limited, virtual space can be designed as large as desired. To facilitate the exploration of the entire virtual space, techniques must be provided to compensate for physical movement. Therefore, the second part investigates potential improvements in orientation by providing additional tools to support visual orientation when the most popular motion compensation technique – i.e., virtual teleportation – is used. This part was published as an evaluation paper at ISMAR [205] and comprises a user study comparing, among other things, performance, efficiency, memorability, and movement behavior of users in different search and exploration tasks when equipped with different visual aids.

✦

4.1 The Impact of Immersion on Cluster Identification Tasks

Recent developments in technology encourage the use of HMDs as a medium to explore visualizations in VR. VREs enable new, more immersive visualization design spaces compared to traditional computer screens. Previous studies in different domains, such as medicine, psychology, and geology, report a positive effect of immersion, e.g., on learning performance or phobia treatment effectiveness. Our work presented in this chapter assesses the applicability of those findings to a common task from the InfoVis domain. We conducted a quantitative user study to investigate the impact of immersion on cluster identification tasks in scatterplot visualizations. The main experiment was carried out with 18 participants in a within-subjects setting using four different visualizations, (1) a 2D scatterplot matrix on a screen, (2) a 3D scatterplot on a screen, (3) a 3D scatterplot miniature in a VRE and (4) a fully immersive 3D scatterplot in a VRE. The four visualization design spaces vary in their level of immersion as shown in a supplementary study. The results of our main study indicate that task performance differs between the investigated visualization design spaces in terms of accuracy, efficiency, memorability, sense of orientation, and user preference. In particular, the 2D visualization on the screen performed worse compared to the 3D visualizations with regard to the measured variables. The study shows that an increased level of immersion can be a substantial benefit in the context of 3D data and cluster detection.

4.1.1 Introduction

Different visualization design spaces, i.e., spaces in which a visualization is projected, exist. Visualizations often need to adapt to the given design space, which can change their level of immersion. An example of a common visualization design space is a two-dimensional space on a monitor screen. Any visualization that encodes a maximum of two attributes with one position can be displayed within this space (e.g., 2D scatterplot [65] or 2D parallel coordinates [169, 401]). Another visualization design space is created if an additional third attribute is encoded in the visual variable “position” (e.g., space time cubes [11, 196] or 3D scatterplots [193, 286, 310]). The level of immersion may already differ

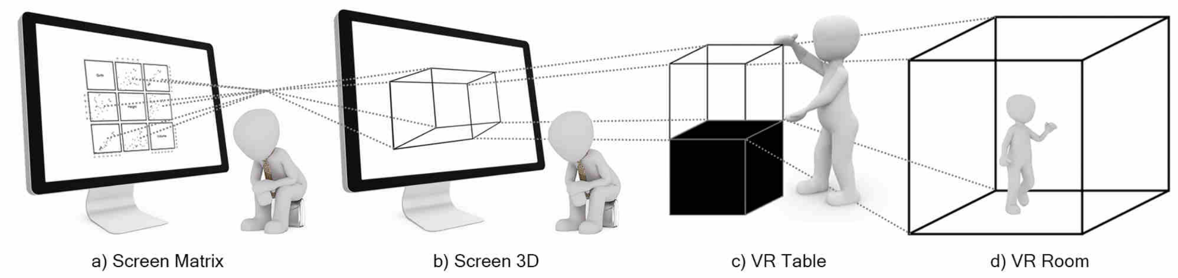

Figure 4.1: A cluster identification task was performed and evaluated in four different visualization design spaces. Two screen-based methods, namely a scatterplot matrix (a) and a 3D scatterplot in a cube (b), and two visualizations in a VR environment: a 3D scatterplot on a virtual table (c) and a room-sized scatterplot (d). Gray lines emphasize transitions between visualization design spaces.

between the two exemplary design spaces (2D and 3D design space) as a higher degree of abstraction is necessary to display the same information in 2D as compared to a more natural display in 3D. For instance, a 3D scatterplot can be visualized in the 2D visualization space as a scatterplot matrix or after a PCA projection in a 2D scatterplot, both being more abstract than a 3D scatterplot visualized in a 3D visualization design space. The more familiar nature of the 3D data representation may lead to a more intense perception of immersion.

Over the last few years, AR, VR, and MR hardware and software have been on the rise, opening up new design spaces for visual analytics applications. Various examples of visualizations exist in VR, AR, and MR, either restricting the visualization's space to a small area [12] or allowing it to occupy the entire space around the observer [111, 213]. As the level of immersion with regard to the visualization differs largely between the two kinds, their visualization design spaces can be seen as two individual ones.

It is often not a trivial decision which design space is best suited for a specific task. There are several studies comparing visualizations in VREs to those in conventional design spaces, but they often focus on differences resulting from different visualization and interaction techniques [12, 388]. These studies do not capture how much of the differences in performance can be ascribed to those two factors and how much to the different levels of immersion. In this chapter, we want to investigate how much influence the choice of the design space, and the associated level of immersion, has on the overall performance of visualizations. Since this is a rather broad question, we specifically focus on the task of cluster detection in scatterplot visualizations. Our study builds upon the work of Wagner Filho et al. [388] who investigated the effects of immersion provided by VREs. However, they compared the level of immersion provided by different interaction techniques and not the level of immersion provided by the design space. In particular, we investigated differences between four visualization design spaces, each having a certain level of immersion. In order to focus on differences due to the design spaces themselves, we minimized user interaction and used three-dimensional data in combination with a simple visualization.

4.1.2 Related Work

In this section, we provide a brief overview of the most related strains of research. First, we target research in which 2D visualizations were deployed and quantified for cluster identification tasks. Second, we present several examples of effects of immersion in various domains, which motivated our research to assess similar effects for InfoVis tasks. At the same time we outline how immersion was measured in previous work. We then focus specifically on 3D visualizations and 3D scatterplots since they are an integral part of the current work for the reason that they serve as a base visualization in the present study. Subsequently, we discuss advantages and disadvantages of data visualizations using stereoscopic displays as the current work investigates possible benefits and drawbacks of visualization design spaces in VR compared to screen-based ones by means of scatterplots.

4.1.2.1 Cluster Identification with static 2D Visualizations

For cluster identification tasks, a number of static 2D visualizations are commonly deployed, such as parallel coordinate plots [157, 418], dendograms [254, 427], and heatmaps [331, 381]. For the analysis of higher dimensional data, scatterplot matrices are a common technique for cluster identificaiton tasks [106, 165].

While various techniques for the visual exploration of previously extracted clusters in scatterplots exist [165, 174], the technique is also deployed for visual identification tasks of clusters. Cavallo et al. [54] propose a framework in which they make use of scatterplots to identify clusters. In their framework, they also deploy other techniques, such as silhouette plots and heatmaps. Etemadpour et al. [112] deployed an eye tracker to monitor user behavior when exploring 2D scatterplots for various tasks. They found that cluster density is more influential than cluster size in cluster identification tasks. They also discuss issues of cluster separation and cluster preservation for deployed dimensionality reduction techniques and their impact on user performance. Therefore, we included cluster density as an experimental side factor.

4.1.2.2 The Effects of Immersion

Several studies have shown that immersion can have a benefit in different fields, e.g., geology, architecture, and medicine. Examples from these areas show that increased immersion can foster spatial understanding and orientation [334] and increase focus capabilities of users by helping them to fade out distractions [25]. It was also shown that a higher level of immersion can increase task efficiency and effectiveness for spatial problem solving applications [141] and psychological treatment procedures (e.g., phobia treatment)[390]. Positive effects of immersion have also been reported on learning performance in the context of medical education [145], on memorization [300], as well as for visualizing abstract visualizations [215, 301].

In many studies it is just presumed, without elaboration, proof or reference, that VREs convey higher levels of immersion than screen-based media. However, to prove this assumption, some metric needs to be introduced measuring immersion. According to Slater et al.[347], immersion can be seen as a rather objective property of a system that introduces a subjective impression of presence to the user. Various researchers intended to measure immersion by quantifying system properties, such as resolution, field of view, degrees of freedom in movement and so on [153, 348]. This is, however, quite hard to quantify and measure. Witmer and Singer [413] propose to measure subjectively perceived presence and infer results back to immersion. In the past, researchers developed and applied several presence questionnaires under that premise [223, 308]. The most established one is the Presence Questionnaire (PQ) from Witmer and Singer [413].

4.1.2.3 3D Visualizations

Previous work has shown that 3D visualizations are often vulnerable to artifacts caused by the rendering of depth-related information, such as line-of-sight ambiguities, occlusion, and perspective distortion [137, 286]. Depending on the viewpoint, the visualization is distorted differently, hampering cognition and impeding the interpretation of distances and proportions between objects. Therefore, visual variables that perform well in 2D visualization spaces, such as length, size or position, may be less suitable in 3D visualizations due to depth distortion and the missing alignment with respect to a common baseline. Nevertheless, there are several advantages of 3D visualizations in general and, hence, various 3D visualization applications exist in different domains [262, 365]. Multiple studies show benefits of 3D in exemplary visualizations [137, 266], among others with regard to accuracy and efficiency. Moreover, studies indicate that 3D visualizations perform even better when inspected using stereoscopic displays due to a more natural, familiar and accurate perception of information [111].

3D scatterplots are used in various applications for visualizing multi-dimensional data [193, 286, 310, 426], e.g., to visualize network data [365] or a development over time in space time cubes [129, 262]. Sedlmair et al. [335] compared 2D scatterplots, 3D scatterplots and scatterplot matrices. They examined the effectiveness of these visualizations for separating clusters in datasets that have been transformed with the help of a dimension reduction technique. They found that 2D scatterplots could be used to perform the given task to a satisfactory extent, but that in most cases participants using scatterplot matrices outperformed others using 2D scatterplots. According to them, using 3D scatterplots for the examined task rarely helped, and sometimes even impaired the results. However, they only used data which previously was subject to a dimension reduction procedure and did not evaluate scatterplots in a VRE. Since they assumed that differences in performance between the designs mainly result from the data and not the users, the study was conducted with only two expert users. Each of them inspected and classified 816 scatterplots.

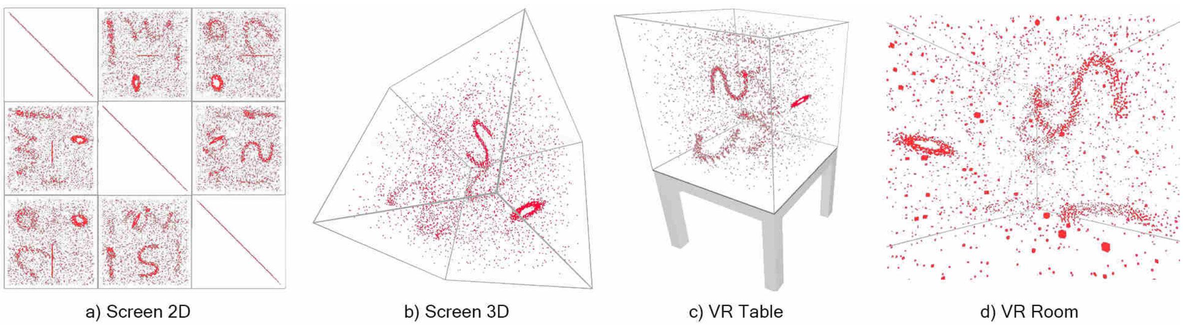

Figure 4.2: Representation of one exemplary dataset in all four investigated visualization design spaces. Except for the scatterplot matrix (a), all visualization design spaces had some kind of navigation available to inspect the visualization from different perspectives.

4.1.2.4 Evaluation of Stereoscopic Visualizations

Wagner Filho et al. compared 2D scatterplots with screen-based 3D scatterplots and VR-based 3D scatterplots [388]. Their tasks included finding nearest neighbors, finding the nearest class, identifying class outliers and comparing two classes to each other. Users in this study were faster using the 2D scatterplot and found it slightly more intuitive for the given tasks. On the other hand, participants were slightly more accurate and subjectively more engaged using the VR scatterplot. In a follow-up study, Wagner Filho et al. [386] present and evaluate an analysis environment in which the user is seated and interacting with scatterplot visualizations using gestures. The authors further investigated user capapabilities to evaluate dimension reduced 3D scatterplot visualizations in immersive and screen-based scenarios [387].

Prabhat et al. [293] conducted a study to evaluate environments differing in their level of immersion by means of different data analysis tasks. However, to the best of our knowledge, there is no study evaluating the impact of the degree of immersion in VREs on user performance during scatterplot analysis. By now, research has not extensively assessed the opportunities and disadvantages of design spaces in VR for abstract visualizations in VA tasks.

4.1.3 Design Spaces

In this chapter, we investigate user performance in a cluster identification task by means of scatterplots in four different visualization design spaces (see concept in Figure 4.1 and realization in Figure 4.2). The conducted study solely targets the visual detection of clusters in a dataset visualized as a scatterplot without encoding the cluster membership of data points and compares user performance in different visualization design spaces. In this section, each of the examined design spaces is briefly described. Subsequently, we reason why we chose the presented design spaces. In our basic scientific research approach, we consider three-dimensional data only. In many cases, multi-dimensional datasets can be effectively projected into 3D space using dimension reduction methods (e.g., PCA[192], t-SNE [127]). However, for truly high dimensional data, projections into 2D or 3D space might not be suitable for cluster identification tasks. In cases like that, three dimensions could be compared at a time in small multiple visualizations. We argue that we investigate basic visual perception and the users' capability to identify clusters in three-dimensional datasets. We only deploy the visual variable position and abstain from using additional visual variables (e.g., color, shape) to keep the experiment as simple as possible. Because a maximum of three dimensions can be encoded in 3D visualizations exclusively by position, we focused on the reduction to three dimensions. Consequences of this constraint, in particular with regard to the 2D design space, are discussed in section 4.1.8.

***Screen2D*: 2D on Screen –** The first design space is a 2D space on a monitor screen. To represent three-dimensional data in two-dimensional space, there are at least two intuitive options. One option is to use a dimension reduction technique, map the data to two dimensions, and visualize the resulting projection in a standard two-dimensional scatterplot. We decided against this approach as sometimes clusters vanish in the projection. Figure 4.3 depicts a two-dimensional projection of three-dimensional data displayed in Figure 4.2 after a PCA dimension reduction. The visualization demonstrates that, for some use cases, a PCA transformation can be unsuitable for cluster identification tasks. In the given example, only four out of six clusters are clearly distinguishable in the PCA projection (see Figure 4.3). More advanced dimension reduction techniques, such as t-SNE [127], often require a set of parameters that must be customized for each dataset to result in an optimal representation for the cluster identification task. In the case of our study, this would require additional user interaction and significantly increase interaction efforts for this visualization design space and consequently affect results. Moreover, due to individual adjustments of parameters, results of different participants would not be comparable anymore.

Another option is the display of 3D data in a scatterplot matrix representation, which is often used to visualize multi-dimensional data in various domains [53, 106, 312]. The scatterplot matrix

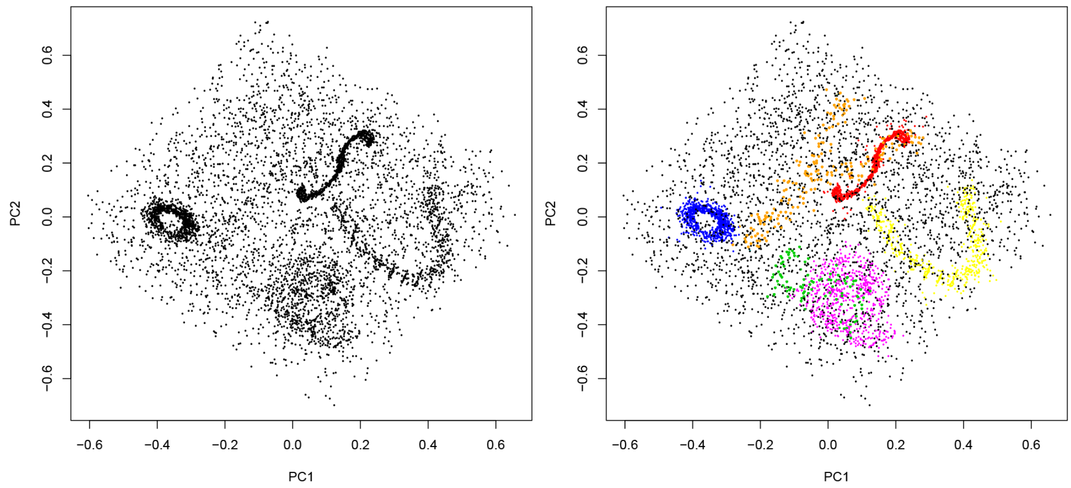

Figure 4.3: PCA projection of data displayed in Figure 4.2. The dataset contains six clusters (highlighted on the right). Two clusters are hardly recognizable in the PCA projection (orange *Y* and green *S*).

is a projection of high-dimensional data into a 2D representation consisting of small multiples (2D scatterplots). For data with three dimensions (x, y, z), the resulting visualization is a compound of three different scatterplots (x&y, x&z, y&z) and rotated and mirrored versions of them as can be seen in Figure 4.2a. In our investigations, we chose to use a scatterplot matrix to visualize the data in this design space. *The observer is looking at a static, non-interactive scatterplot matrix on a screen.*

***Screen3D*: 3D on Screen –** The second design space is a 3D space on a monitor screen. The resulting visualization is a virtual 3D cube on a screen, containing the three-dimensional data as a 3D scatterplot. This design space is also frequently deployed in related works [106, 193, 286]. *The observer is looking at a projection of a 3D visualization on a screen, inspecting the data by rotating the scatterplot in arbitrary directions.*

***VRTable*: Miniature 3D in VRE –** The third design space is a restricted 3D space in a VRE. This design space is limited spatially so that the observer is able to walk around the visualization and observe it from outside. The resulting visualization is a 3D scatterplot on a virtual table in a VRE (table height: 75 cm; cube dimensions: 1 m $\times$ 1 m $\times$ 1 m; data point size: 2.5 mm). *The observer is standing in front of a virtual table with a 3D scatterplot on top of it, inspecting it by walking around the table.*

***VRRoom*: Room-Scaled 3D in VRE –** In the fourth design space, we adjusted the size of the 3D scatterplot to the size of the entire VRE (dimensions: 3 m $\times$ 3 m $\times$ 3 m; data point size: 7.5 mm). The entire space around the observer is used as visualization design space. *The observer is standing inside the visualization and inspects the scatterplot from within by walking and looking around.*

Design Decisions: In order to investigate the effects of immersion provided by design spaces, we aimed to create several different design spaces with varying degrees of immersion. First, we chose to introduce a 2D design space located on a 2D screen (*Screen2D*). This is a commonly used design space and can be seen as a baseline for the other design spaces. In line with the definition of immersion by Slater [345], we perceive a virtual object as more immersive if it reflects the characteristics of a real object. Therefore, to increase the degree of immersion provided by the design space, the resemblance with real-world objects has to be increased for virtual objects. This can be achieved by using a 3D design space located on a 2D screen (*Screen3D*). Thereby, data points are displayed more “naturally” as the real world is 3D itself. Moreover, with regard to scatterplots, we can easily perceive all three dimensions at the same time in a 3D environment, whereas heavy mental mapping is required to extract all dimensions of a data item from a scatterplot matrix.

Presenting a 3D object on a 2D screen usually introduces perspective distortions [137]. These distortions change how a person perceives the object and, therefore, may reduce immersion as the object reflects the characteristics of a real object to a lesser extent. This effect can be avoided by using VREs. Therefore, we deployed VR in the third and fourth design space. In the third design space (*VRTable*), we introduced the restriction that the user can only observe the visualization from outside and is not able to enter the visualization itself. We argue that this restriction is insofar reasonable as the same restriction applies to the previous design spaces. Removing the restriction (*VRRoom*) may lead to an even more increased level of immersion as the user enters the visualization itself and is fully enclosed by it.

In order to validate our hypothesis that the level of immersion increases in each design space (see Figure 4.2, a to d), we conducted a supplemental study in which we investigated solely this specific issue (section 4.1.4).

4.1.4 Pre-Study: Levels of Immersion

Among others, the property *level of immersion* discriminates visualization design spaces. According to Slater [345], immersion describes how much a system preserves the fidelity of sensory modalities. To confirm differences between the proposed design spaces presented in section 4.1.3, with regard to their level of immersion, we conducted a pre-study. As it is hard to directly measure the properties of the system, we rely on the approach of Witmer and Singer [413], i.e., measuring presence and referring it back to immersion. In this pre-study, we evaluated participants' level of self-reported immersion for each design space. Further subjective observations, opinions and perceptions of participants concerning the design spaces (e.g., abstractness, preference) were gathered.

4.1.4.1 Study Description and Hypothesis

The only experimental factor of this study was *visualization design space*. All four design spaces introduced in section 4.1.3 were examined by means of a within-subjects design. We hypothesize that the design spaces can be sorted by their level of immersion as follows: *Screen2D* is the least immersive design space, followed by *Screen3D*, *VRTable* and *VRRoom*. As there was no reason to assume an impact on participants' physical or mental health, no institutional review board (IRB) was consulted for the study. Also, the participants could abort the study at any time.

After a training session in all design spaces, 12 participants (six female, six male) conducted one cluster identification task in each design space. The order of designs and used datasets were counterbalanced. All datasets had similar properties, contained between five and seven clusters and were enriched with the same amount of noise. Subjects were asked to identify and count all clusters in the data and to report their result to the examiner. After each of the four trials, participants completed a questionnaire. Both a multiple measures questionnaire for immersion (consisting of 18 questions) and a single measure of immersion (consisting of one question) were used. The first question served as a single measure of the subjectively perceived immersion in the respective design space: "How immersed did you feel in the virtual environment?". The following set of 18 questions were adopted from questionnaires by Regenbrecht et al. [308] (IPQ), Witmer and Singer [413] (PQ), Lessiter et al. [223] (ITC) and Jennett et al. [177] (IEQ). We carefully selected questions that fit all design spaces as well as the current task. Therefore, we excluded, for instance, questions that are explicitly aimed at gaming experiences in VREs. After the completion of all four trials, we conducted a semi-structured interview. At the end of the experiment, participants received 10 Euros as compensation. The apparatus of this study was similar to the one in the main study described in section 4.1.5.4.

4.1.4.2 Results of Pre-Study

Questionnaires Statistical tests were performed using IBM SPSS Statistics (version 24). In this section, we only report significant results. A Bonferroni correction was applied (to control for multiple testing) and, hence, all effects are reported at a .008 level of significance.

To evaluate differences in the level of immersion between design spaces with regard to the single measure of immersion, a non-parametric Friedman test was deployed ($\chi^2(3) = 22.49$, $p < .001$). We used a non-parametric test because of skewed distributions. Wilcoxon signed-rank tests were computed as post hoc tests to follow up this finding (see Figure 4.4 left). The post hoc tests revealed that the subjective experience of immersion was significantly lower in the *Screen2D* design space (Mdn = 1.00) as well as in the *Screen3D* design space (Mdn = 2.50) compared to both VR spaces, namely *VRTable* (Mdn = 4.00) and *VRRoom* (Mdn = 4.00): *Screen2D-VRTable*: $z = -2.84$, p = .001; *Screen2D-VRRoom*: $z = -2.75$, p = .002; *Screen3D-VRTable*: $z = -2.85$, p = .001; *Screen3D-VRRoom*: $z = -2.61$, p = .003.

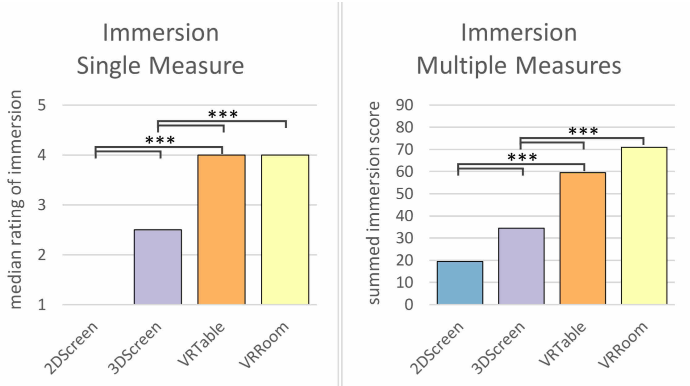

Figure 4.4: *Measures of Immersion* – Left: single measure question of subjectively perceived immersion. Right: multiple measure questionnaire on immersion.

For the multiple measure of immersion (i. e. the immersion questionnaire), immersion scores were computed by summing up participants' responses to all 18 questions. The same statistical approach was used as for the analysis of the single measure of immersion ($\chi^2(3) = 24.23$, $p < .001$). As depicted in Figure 4.4 (right), Wilcoxon signed-rank tests showed that the level of immersion was significantly lower in both the *Screen2D* design space (Mdn = 19.50) and the *Screen2D* design space (Mdn = 34.50) than in the two VR spaces, namely *VRTable* (Mdn = 59.50) and *VRRoom* (Mdn = 71.00): *Screen2D-VRTable*: $z = -2.90$, p = .001; *Screen2D-VRRoom*: $z = -2.90$, p = .001; *Screen3D-VRTable*: $z = -2.94$, $p < .001$; *Screen3D-VRRoom*: $z = -2.87$, p = .001.

Interview The evaluation of the interview questions on abstraction and presence, which can be regarded as substitute variables for immersion [413, 419], revealed the predicted order of design spaces (see median user ratings depicted in Figure 4.5, left and center)). Participants perceived the VR design spaces as less abstract and therefore more natural compared to the two screen-based ones

(*Screen2D*: Mdn = 5; *Screen3D*: Mdn = 2.5; *VRTable*: Mdn = 2; *VRRoom*: Mdn = 2). Particularly for the subjective user rating of how present they felt in the respective design space, the assumed pattern emerged (*Screen2D*: Mdn = 1; *Screen3D*: Mdn = 2; *VRTable*: Mdn = 4; *VRRoom*: Mdn = 5).

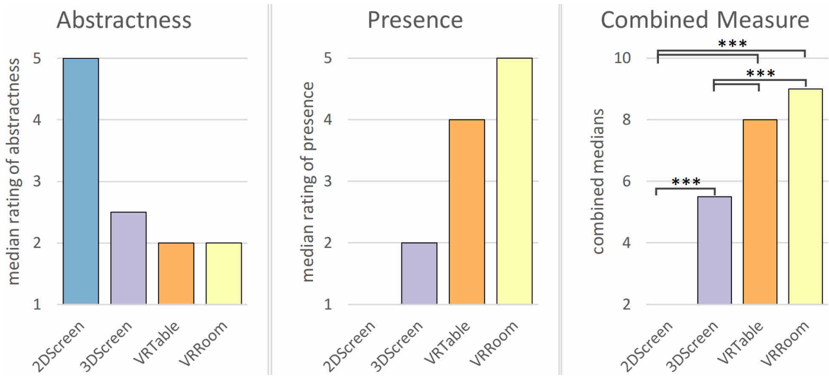

Figure 4.5: *Interview* – Median user ratings for the design spaces with regard to abstractness (left) and presence (center). Participants were asked to rate the abstractness and presence of each design space on a five-point Likert scale from 1 = *not abstract/not present* to 5 = *very abstract/very present*. Right: Combined median of the abstractness and presence scores used for statistical evaluation and as a measure of immersion.

We conducted a Friedman test ($\chi^2(3) = 30.28$, $p < .001$). Bonferroni-corrected Wilcoxon signed-rank tests revealed significant differences between all design spaces: *Screen2D-Screen3D* ($z = -2.89$, p = .001), *Screen2D-VRTable* ($z = -3.27$, $p < .001$), *Screen2D-VRRoom* ($z = -3.28$, $p < .001$), *Screen3D-VRTable* ($z = -2.75$, p = .002), *Screen3D-VRRoom* ($z = -2.97$, p = .001), *VRTable-VRRoom* ($z = -2.67$, p = .005).

These results are supported by the interview question, in which the participants were asked to sort the design spaces by the amount of perceived presence. All subjects put *VRRoom* in first place (n = 12) and *Screen2D* last. Only one participant put *Screen3D* in second place and *VRTable* in third place – all others put the design spaces in the expected order.

4.1.4.3 Conclusion

Overall, our pre-study supports the previously stated hypothesis and verifies the assumed order of design spaces with regard to the level of immersion:

$$Screen2D < Screen3D < VRTable < VRRoom$$

4.1.5 Main Experiment

As shown in previous research, the degree of immersion can have an effect on spatial cognition and memorability in various contexts [80, 229]. Some studies even indicate correlations between the degree of immersion and efficiency in cluster identification, distance estimation, and outlier detection tasks in scatterplot visualizations [12, 301]. However, many existing studies use a variety of different interaction techniques individually for each design space, disguising possible effects caused solely by characteristics of the different design spaces. In order to avoid possible confounding factors resulting from different interaction techniques, we limited our study to an absolute minimum of interaction techniques. No institutional review board (IRB) was consulted for the study as there was no reason to assume any impact on participants' physical or mental health. Participants could end the study at any point.

4.1.5.1 Study Design

Our main experimental factor was the *visualization design space*. Besides, we investigated the impact of noise level, cluster shape, and cluster density. All three study side factors (noise, shape, density) were introduced to examine if the designs are differently robust to dataset characteristics and to ensure that our main results are generalizable to different kinds of datasets. A prototype, developed specifically for the purpose of this study, was used for the execution of the study.

Visualization Design Space: As main experimental factor the design spaces introduced in section 4.1.3 were examined. In each design space, an adaption of a scatterplot was displayed (scatterplot matrix, 3D scatterplot). Figure 4.2 shows one exemplary dataset in all four designs.

Noise Level: The first additional experimental factor was the level of introduced noise. With regard to the noise level, two kinds of datasets were generated. One contained 1,000 additional randomly positioned points (low noise level), the other one 3,000 additional noise points (high noise level).

Cluster Shape: As a second additional experimental factor, the shape of clusters was manipulated. Half of the datasets contained convex clusters (spheres, capsules, discs), the other half contained non-convex clusters (spirals, donuts, y-shapes, s-shapes, sinus-curved pipes). We statistically counterbalanced noise level and cluster shape, i.e., all four possible combinations (low-noise & convex, low-noise & non-convex, high-noise & convex, and high-noise & non-convex) occurred equally often.

Cluster Density: The third additional experimental factor was cluster density. Two different types of clusters were created with regard to cluster density. We used the DBSCAN algorithm, introduced by Ester et al. [110], as a measure to distinguish between dense and sparse clusters. For sparse clusters the parameters $t_1 = \{ MinPts = 10, \varepsilon = 0.15 \text{ m} \}$ were used as thresholds, and for dense clusters $t_2 = \{ MinPts = 30, \varepsilon = 0.10 \text{ m} \}$. The two parameter sets were systematically refined during several trial dataset generation procedures to generate two visually distinguishable types of clusters. The clustering was performed in a cube with the dimensions 2 m $\times$ 2 m $\times$ 2 m. With the lower threshold t_1, all clusters, but nothing else, should be found by the DBSCAN algorithm, and with the higher threshold t_2, solely all dense clusters should be found. In contrast to the other three side factors, each dataset contained both sparse and dense clusters at the same time. However, the error rate was measured separately for both types of clusters.

4.1.5.2 Procedure

The experiment was structured in four blocks. Each block was dedicated to one visualization design space (*Screen2D*, *Screen3D*, *VRTable*, *VRRoom*). In each block, the participant completed four trials by pointing to all clusters found and reporting the overall count to the study supervisor. Each trial had a different dataset. The order of blocks was structurally alternated with the only constraint that always the two screen-based and the two VR design spaces were directly after each other. We chose to introduce this constraint because pretests showed that some participants experienced varying levels of discomfort after switching in or out of the VRE. Half of the participants started with VR design spaces, half of them with screen-based ones. Participants were systematically assigned to one order.

At the beginning of the experiment, written informed consent was obtained from the participants and they were asked to fill in a questionnaire assessing demographic variables. After that, participants completed four blocks, each beginning with a training session for the respective visualization. A total of three practice trials had to be completed before the first real trial of the block could start. In each trial block, participants completed eight tasks. At the end of the second block, participants were again asked to fill in a brief questionnaire examining participants' memory of the last completed trial. A third questionnaire was administered after the last block, collecting information about personal preferences and subjective opinions about the four visualizations. Finally, participants were thanked and received a monetary compensation for participating (10 Euros). During the experiment, sound, video, and position data were recorded.

4.1.5.3 Data and Task

Sedlmair et al. [336] proposed a taxonomy of visual cluster separability factors in scatterplots. They describe various factors of clusters, such as shape, size, or number of items, that affect the observer's capability to identify the centroid of each cluster in dimensional reduced datasets presented as 2D scatterplots. For the generation of our study datasets, we varied the identified variables shape, size and density. We first created a set of 16 different clusters. In this context, we refer to clusters as areas with a higher density of data points compared to the surrounding areas. In order to guarantee the cluster property and also a consistency over all clusters, we applied the DBSCAN algorithm after creating the clusters interactively in Unity. Two different types of clusters were prepared with regard to density. Dense and sparse clusters had to be found as only clusters by the DBSCAN algorithm with a certain parameter set ($MinPts = 10$, $\varepsilon = 0.15$). Sparse clusters must not be found using another parameter set ($MinPts = 30$, $\varepsilon = 0.1$), which should only detect all dense clusters (see section 4.1.5.1).

Finally, we generated 32 study datasets as compositions of rotated and flipped versions of the previously created clusters. Additionally, we created a set of 20 extra datasets for training trials. Subsequently, we added a certain amount of noise to each dataset (50% of the datasets with high noise level). Half of the datasets contained only convex clusters, and the other half only non-convex shaped ones. Each dataset was constructed carefully so that all clusters were potentially identifiable in the scatterplot matrix (i.e., no cluster was occluded in all views). Exemplary datasets from both conditions are depicted in Figure 4.6.

Although all our datasets were created with three attributes (one coordinate each for the *x-axis*, *y-axis*, and *z-axis*), we do not see this as a limitation of our study. Higher dimensional data can be transformed into 3D data by projection techniques like a PCA. However, the type of projection and its settings has a major impact on how well clusters can be identified in the resulting visualization. For this study, we created datasets natively in three dimensions and abstained from deploying dimension reduction techniques as it is common practice in real-world applications. We only aim to investigate the effects of immersion provided by the design spaces, which should not be affected by a preceding data transformation step.

For the entire experiment, the task performed by participants remained the same, even though interaction and visualization techniques differed. The task was to identify clusters in a scatterplot visualization, to point at them, and to count up all clusters. Participants were asked to point at found clusters (with the mouse or VR controller) and report their detection to the study supervisor. At the end of each trial, they indicated the overall count of found clusters.

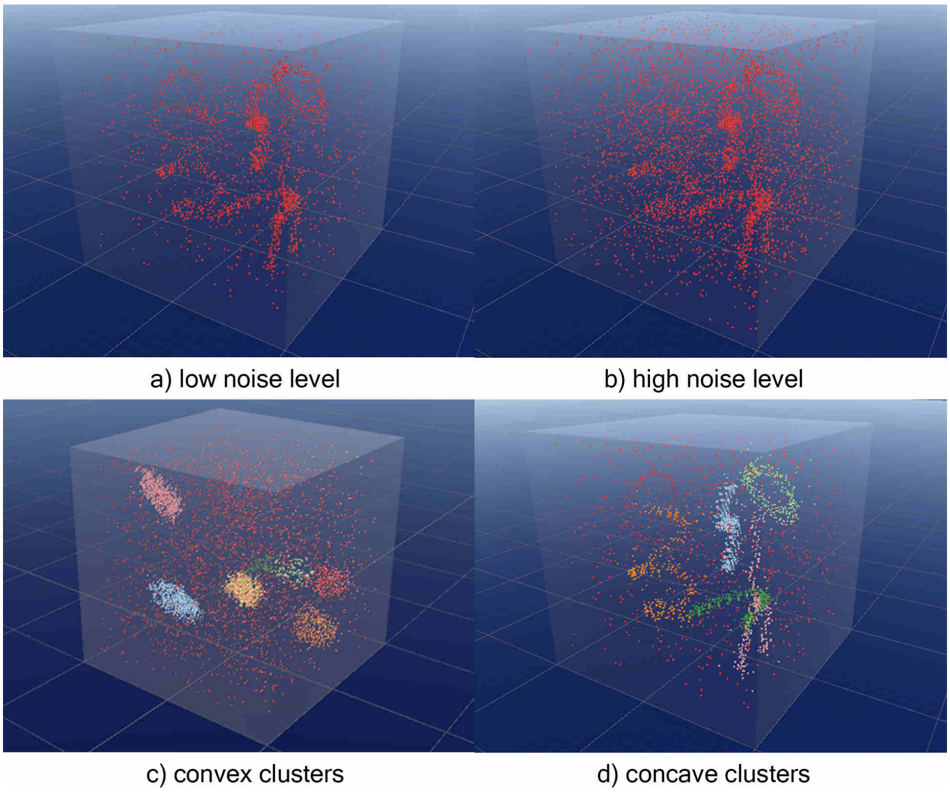

Figure 4.6: Four sample datasets illustrating different properties. Top: low noise condition (left) and high noise condition (right). Bottom: convex clusters condition (left) and non-convex clusters condition (right).

4.1.5.4 Apparatus

The experiment took place in a quiet, closed room at the University of Konstanz. Participants were individually invited to the laboratory. Besides the participant, the examiner was the only person present. During two blocks (screen-based visualization design spaces), the participants sat in front of a 24" monitor with a resolution of 1920 $\times$ 1200 pixel. In those blocks, participants interacted with the study software solely with the mouse as input device. During the remaining two blocks, participants were equipped with a Vive HMD and one Vive controller as a pointer. In those blocks, participants were initially positioned at a specific starting point. During the task, they were allowed to walk freely through the room within the bounds of the virtual environment (which were visually highlighted in the VRE as blue walls). In the *VRTable* visualization design space, participants were additionally instructed not to walk into the virtual table.

4.1.5.5 Sample

A sample of $N = 18$ participants (5 female, 13 male) was recruited using short notices distributed around the university. Most of the participants had none or only little experience with scatterplot matrices (66.6%), but had experienced a VRE at least once before (72.2%). We introduced a training phase at the beginning of each of the four blocks in order to minimize any effects resulting from different levels of experience. Participants were aged 19 to 41 years ($M = 26, SD = 4.87$). Three participants were still in high school, nine held a Bachelor's degree, and six a Master's degree. The background of the participants was quite diverse with eight having a computer science background (44%) and the rest from various domains without advanced computer science knowledge.

4.1.5.6 Dependent Variables

To compare differences caused by changes in the independent variables (visualization design space, noise level, cluster shape, and cluster density), we analyzed multiple dependent variables. For each trial, the error rate was calculated as the percentage of clusters not found. All trials were recorded for later video analysis to count errors and find frequent patterns in participant behavior. Participants were instructed to point at identified clusters throughout the trials using the mouse (screen) or the laser pointer attached to the Vive controller (VR). After the study, we analyzed the recordings by coding which clusters were found in each trial. All videos were encoded by at least two people to avoid counting errors. In addition, the task completion time was logged. For all VR trials, the VR headset was tracked (head position and orientation). Besides, two questionnaires were issued gathering information about personal preferences and the memorability of data (count, shapes, and positions of clusters) in a previously completed task.

4.1.5.7 Hypotheses

Based on subjective indications from two exploratory pilot studies and in part based on results of studies presented in the related work section, we derived the following hypotheses. All hypotheses refer to the deployed tasks and variants of scatterplot visualizations.

H1 VR vs. Screen – Error Rate: We expect error rates to be lower when participants work with VR visualization design spaces. This hypothesis is based on Filho et al. [387] and Arns et al. [6] experiments on the analysis of multi-dimensional data in 3D scatterplots. They report on beneficial effects of immersion with regard to distance and structure perception. Both properties are crucial for cluster identification.

H2 VR vs. Screen – Task Completion Time: Participants will be more active and need more time to complete the task when working in VR visualization design spaces compared to them working in screen-based ones. Bach et al. [12] came to the conclusion that participants need more time in AR environments because they move more, take extra time to explore the visualization, and are new to the device. We expect similar findings in our VR settings.

H3 VR vs. Screen – Memorability: Participants will show better memory performance when working in VR design spaces compared to them working in screen-based design spaces. Previous studies have shown that in certain VR scenarios the spatial memory is crucially better compared to applications on the screen due to a more natural navigation [67].

H4 VR vs. Screen – Subjective Preference: Visualizations in VR visualization design spaces will come more naturally to the participants than the ones in screen-based design spaces. This hypothesis is based on the assumption that the level of abstraction of VR visualizations should be relatively small as, for instance, distances can be measured in “real” measures such as inches or centimeters. Additionally, the possibility to navigate the data space like in the real-world (e.g., walking around or rotating the head) is expected to increase the engagement of participants [12].

H5 Full Environment vs. Restricted Area – Error Rate: Comparing the VR visualization design spaces, participants will perform worse in the totally immersive design space (*VRRoom*) compared to the *VRTable* design space with regard to the error rate. This hypothesis is based on the assumption that participants will miss clusters due to blind spots (clusters behind, underneath or above the observer) or a possible loss of orientation due to the missing overview as reported by Etempadpour et. al. [111].

4.1.6 Results of Main Study

We report significant results of our quantitative analysis, as well as qualitative feedback.

4.1.6.1 Statistical Analysis

All statistical tests were performed using IBM SPSS Statistics (version 24) and are based on a significance level of $\alpha = .05$. To evaluate differences between the visualization design spaces related to the error rate, i.e., the percentage of clusters not identified, a Friedman test was used. Due to serious violations of assumptions, in this case we have decided against an ANOVA and opted for its non-parametric counterpart. Wilcoxon signed-rank tests were computed as post hoc tests. Moreover, a one-way repeated measures ANOVA was applied to compare the time participants required for performing the task (completion time). Mauchly's sphericity test was used to confirm the sphericity assumption needed for a one-way repeated measures ANOVA.

In case of a significant omnibus F-test, we report the results of Bonferroni-corrected pairwise comparisons. Finally, head rotation data were analyzed using a paired samples t-test. Note that time data and head rotation data were log-transformed because of skewed distributions. Shapiro-Wilk tests were used to check the assumption of normality after the log transformations and before the t-tests.

4.1.6.2 Error Rate

Error rates differed significantly between the visualization design spaces ($\chi^2(3) = 40.67$, $p < .001$). As depicted in Figure 4.7 (left), Wilcoxon signed-rank tests revealed that with regard to the error rate participants performed significantly worse in the design space *Screen2D* ($Mdn = 16.67\%$) compared to all other design spaces: *VRTable* ($Mdn = 0\%$, $z = -4.87$, $p < .001$), *VRRoom* ($Mdn = 0\%$, $z = -4.57$, $p < .001$) and *Screen3D* ($Mdn = 14.29\%$, $z = -3.95$, $p < .001$).

When also taking noise into account, there was a significant difference in error rates between the low noise ($Mdn = 8.45\%$) and the high noise condition ($Mdn = 12.47\%$; $z = -2.24$, $p < .025$). For each visualization design space, error rates increased with an increasing noise level. However, the resulting change differed between the design spaces: The difference in error rates between the low noise and the high noise condition was 6.54% for *Screen2D*, 1.3% for *Screen3D*, 5.74% for *VRTable* and 5.54% for *VRRoom*. Statistical tests showed significant differences between noise conditions in both VR design spaces (*VRTable*: $t(17) = -2.27$, $p < .05$, $r^2 = .23$; *VRRoom*: $t(17) = -2.19$, $p < .05$, $r^2 = .22$).

With regard to the side experimental factors cluster shape and cluster density, no significant differences emerged with respect to the error rate.

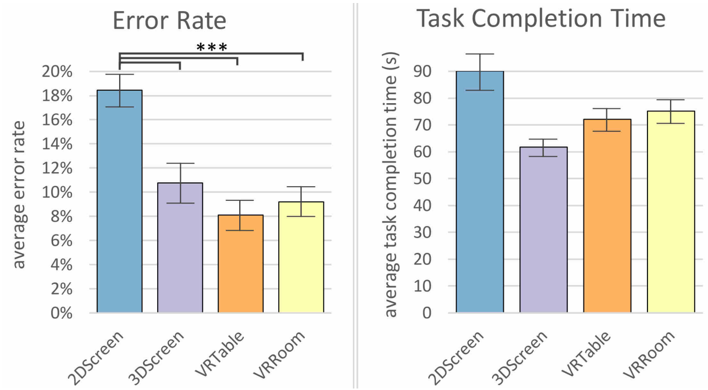

Figure 4.7: Average error rate and completion time as a function of visualization design space. Bars indicate the 95% CI of the mean, asterisks significant differences between design spaces (*** $p \leq .001$). Note that for statistical analysis task completion times were log transformed because of skewed distributions, while in this figure original data is displayed.

4.1.6.3 Task Completion Time

As depicted in Figure 4.7 (right), the average completion time in the *Screen3D* design space ($M = 61.79$ s) was the lowest, followed by *VRTable* ($M = 72.14$ s), *VRRoom* ($M = 75.19$ s) and *Screen2D* ($M = 90.12$ s). Task completion times differed significantly between the four design spaces, $F(3,51) = 4.4$, $p < .01$, $\eta_p^2 = .206$. Bonferroni-corrected post hoc tests were applied. After correcting for alpha error accumulation, none of the pairwise comparisons reached significance. The experimental side factor noise level had no significant influence on task completion time.

4.1.6.4 Memorability

In the questionnaire which was administered after the second block participants were asked to recall the count, shapes, and positions of all clusters in the last completed trial. Results show that, with regard to the error rate, participants performed better in the *VRTable* design space ($M = 0\%$) compared to all other design spaces (*Screen2D*: $M = 43.33\%$; *Screen3D*: $M = 32.67\%$; *VRRoom*: $M = 20.42\%$). The percentages reflect how many clusters of the previously found clusters could not be remembered with the correct shape. To prevent training effects, each participant performed the memory task only once. Therefore, the sample size per design space is rather small ($n \approx 4$).

4.1.6.5 Subjective Preference

As part of the final questionnaire, participants were asked to rank the visualizations by difficulty (1 = *easy* to 4 = *hard*). As Figure 4.8 depicts, ranks assigned to visualizations in the *Screen3D* design space show a positive skewness (ranks 1 & 2: 61.1%; ranks 3 & 4: 39.9%). To visualizations in the *Screen2D* design space, participants only assigned the lowest ranks 3 (22.2%) and 4 (77.8%). To visualizations in the *VRRoom* design space, mainly middle ranks were assigned (rank 1: 11.1%; ranks 2+3: 88.9%). The distribution of visualizations in the *VRTable* design space is positively skewed with the mass center on the upper ranks (ranks 1 & 2: 83.3%; ranks 3 & 4: 16.7%).

In accordance with these results, 50% of the participants mentioned the *VRTable* design space as their preferred design space, 33.3% the *Screen3D* design space and 16.7% the *VRRoom* design space. In contrast, none of them indicated the *Screen2D* design space as their preferred visualization design space. Regarding disadvantages and opportunities perceived by participants, several findings emerged. As

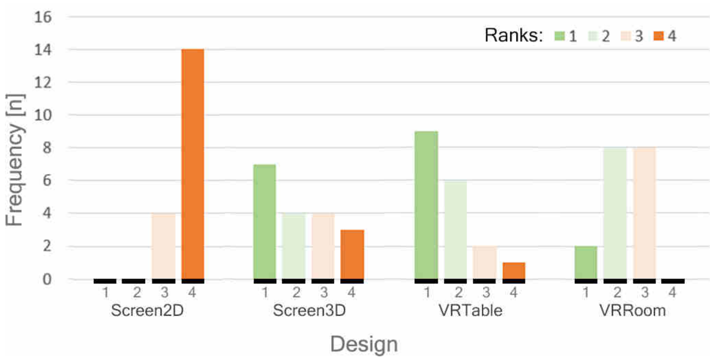

Figure 4.8: Subjective preference: ranks (1 = *easy* to 4 = *hard*) assigned to the four visualization design spaces by the participants.

benefits of VR visualizations, participants rated VR design spaces to be more comprehensive ($n = 8$), intuitive ($n = 5$) and to provide a better overview when the visualization is inspected from outside (*VRTable*, $n = 2$). Moreover, participants mentioned that naturally changing the perspective (moving the head) helps to grasp the visualization ($n = 4$). As drawbacks, participants mentioned poor overview in the *VRRoom* environment ($n = 9$), increased expenditure of time ($n = 4$) and expensive hardware ($n = 2$).

4.1.6.6 Space Utilization and Motion

Except for the area in which the table was located (i.e., the area participants were instructed not to cross), participants used approximately the same amount of space in the two VR visualization design spaces (see Figure 4.9). However, total walking distances varied significantly between the design spaces, $t(17) = -8.80$, $p < .001$, $r^2 = .82$. In the fully immersive environment (*VRRoom*), participants covered considerably more distance ($M = 32.20$ m, $SD = 0.14$) compared to the less immersive design space (*VRTable*: $M = 16.77$ m, $SD = 0.18$).

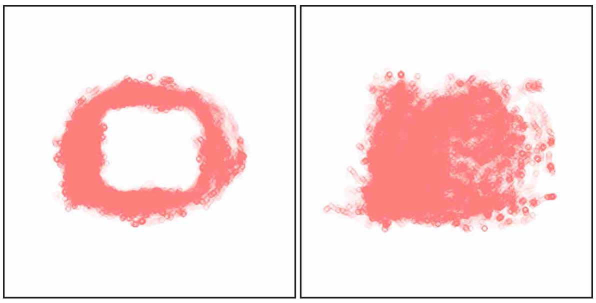

Figure 4.9: Top-down view on the VRE. Participants' movements while solving study trials in the design space *VRTable* (left) and *VRRoom* (right). In the *VRTable* environment, participants were explicitly asked not to walk into or through the virtual table. Except for this, the area covered is approximately the same. However, in the *VRRoom* environment, participants covered roughly twice as much distance compared to the *VRTable* environment.

The two VR environments varied significantly concerning the head rotations of participants, $t(17) = -8.80$, $p < .001$, $r^2 = .82$. In the fully immersive environment (*VRRoom*), participants tended to look around much more ($M = 32196°$, $SD = 0.03$) compared to the less immersive design space (*VRTable*: $M = 16771°$, $SD = 0.04$).

4.1.6.7 Video Analysis

In order to evaluate the experimental trials, we manually examined the videos of each trial. While watching the videos, not only participants' final answer, but also observations throughout the entire task were noted down in a database. We identified several mistakes that were made repeatedly by participants. In particular, four frequent scenarios could be observed: (1) the participant "finds" a cluster twice (double count), (2) the participant finds all clusters, but skips one in the final counting, (3) the participant counts a sparse and a dense cluster as one and (4) the participant detects a sparse cluster, but neglects it as noise. For each of these scenarios, we manually counted the number of occurrences.

The final comparison revealed that all double count-errors were made in screen-based design spaces (*Screen2D*: $n = 7$, *Screen3D*: $n = 2$). Moreover, most adding up-errors (missing to count a cluster in the end) appeared in screen-based design spaces (*Screen2D*: $n = 7$, *Screen3D*: $n = 4$, *VRTable*: $n = 2$, *VRRoom*: $n = 3$). Only in the *Screen2D* design space, it occurred that participants counted a sparse and a dense cluster as one ($n = 2$). Mainly in VR-based design spaces, participants tended to neglect detected sparse clusters as noise (*Screen2D*: $n = 1$, *Screen3D*: $n = 2$, *VRTable*: $n = 1$, *VRRoom*: $n = 5$).

4.1.7 Discussion

In this section, the results, as well as their implications, are discussed with the focus being on accuracy, efficiency, memory, and orientation. In the course of the discussion, we will address all hypotheses.

4.1.7.1 Accuracy

H1 implies that in cluster identification tasks error rates are directly influenced by the degree of immersion present in the respective design space when comparing screen-based design spaces with VR-based ones. Specifically, we assumed that participants perform better in design spaces characterized by higher immersion levels (VR-based design spaces). The results partially correspond to that assumption. In case of the *Screen2D* design space, significant differences emerged. As a basis for this hypothesis it was suggested, among other things, that VR visualizations come more naturally to participants in comparison with abstract visualizations or non-stereoscopic 3D visualizations on the screen (cf. H4). Video analysis revealed that situations containing a loss of orientation or navigational problems mostly occurred in the screen-based design spaces. This indicates improved navigation and orientation capabilities in VREs, which again could be due to a better spatial memory (see section 4.1.7.3). Hypothesis H4 is also strongly supported by qualitative feedback. Multiple participants stated to prefer VR visualizations due to a more comprehensive and intuitive representation of the data. Moreover, participants tended to classify VR visualizations as rather easy to work with compared to screen-based ones (in particular, the scatterplot matrix was frequently rated as the most difficult visualization). For the two VR visualization design spaces differed, no significant difference was found in terms of accuracy. Hence, we cannot confirm hypothesis H5.

4.1.7.2 Efficiency

Contradicting hypothesis H2, no significant differences emerged between pairwise-compared visualization design spaces in terms of task completion time. Nevertheless, the hypothesis can partially be accepted as the statistical analysis revealed a main effect of visualization design space on task completion time and an almost significant difference between the design spaces *Screen3D* and *VRRoom*.

Moreover, the average completion time in the design space *Screen3D* was lower compared to the average completion time in both VR design spaces. One reason for that could be the requirement for the user to be more active in VR design spaces. Instead of sitting in front of a computer screen and operating a mouse, the participant had to move and look around. Tasks in the *Screen2D* scenario required on average much more time than in all other design spaces. This could partially be due to a high learning curve for scatterplot matrices due to small multiples. Participants had to mentally match data points in different visualizations in order to avoid counting a cluster twice or missing one. However, the evaluation of participants familiar with scatterplot matrices did, as well, not reveal a difference. Corroborating the second part of the hypothesis (activity of participants), the total walking distance and the total head rotation differed significantly. The means of both attributes are on average approximately twice as large in the *VRRoom* design space. Possible reasons can be derived from video analysis and user feedback. Participants had to change their position more often in the *VRRoom* design space in order to prevent occlusion or blind spots and they had to turn their heads 360 degrees in order to observe the entire visualization space. One trade-off of the “natural” navigation in VR design spaces is the necessary activity compared to conventional media. Especially for long sessions, the increased physical effort could lead to fatigue, which in turn could affect accuracy and efficiency. Therefore, if using VR design spaces, present findings suggest favoring the *VRTable* design space as it minimizes the required physical activity.

4.1.7.3 Memory and Orientation

Participants performed better with regard to memorizing previously identified clusters in VR visualization design spaces compared to screen-based ones. In the *VRTable* scenario, participants had the least difficulties remembering all clusters and their shapes correctly. Moreover, video analysis revealed that more memory-related errors, such as double counts or missing counts, occurred within screen-based design spaces. Therefore, H3, which states an advantage of VR visualization design spaces in terms of memorability, can be considered confirmed.

After working with the abstract visualization (scatterplot matrix), participants had most difficulties to recall all found clusters. We assume that the higher level of abstraction compromises users’ orientation capabilities, as building a mental model of the small multiples is necessary to notice connections between clusters in different windows of the scatterplot matrix (e.g., to find one cluster in all views). An increased level of difficulty, accompanied by the requirement for a mental model, is also evident from user feedback. Participants voted the *Screen2D* design space to be the most difficult and least preferred design space.

4.1.8 Limitations and Generalizability

Some limitations need to be taken into account. It is discussable whether and to what extent our findings are generalizable and transferable to other visualizations in the given visualization design spaces. We argue that most of the findings rather refer to properties of the design spaces than to characteristics of the individual visualizations (e.g., immersion, spatial memory, orientation, or navigation). Nevertheless, it has to be investigated if found distinctions between design spaces also emerge if alternative visualizations are employed. Changing the type of visualization or allowing more advanced interaction techniques might redistribute assigned characteristics to the visualization design spaces and influence final outcomes.

The *Screen2D* visualization design space is fundamentally different from the other design spaces hampering pairwise comparisons. The scatterplot visualization in the *Screen2D* design space is fixed to a certain viewpoint and does not provide any interactions aside from pointing on clusters. During the generation of datasets, we made sure that every single cluster is potentially detectable in the scatterplot matrix visualization as well and avoided pairs of clusters that overlap in all small multiples of the matrix. Additionally, the data used for the experiment was three-dimensional. In the all 3D design spaces (*Screen3D*,*VRTable* and *VRRoom*) the data was visualized in its natural space whereas in the *Screen2D* scenario, multiple 2D scatterplots had to be displayed to compensate for the third dimension.

Another limitation of the present study is the exclusive deployment of a cluster identification task. Compared to the *VRTable* design space, the *VRRoom* design space helps to reduce occlusion since the visualization occupies the entire virtual environment of the observer. However, this comes at the price of tremendous overview loss. These properties likely have a different impact on cluster identification tasks compared to other visual analytics tasks. Future studies should investigate whether a combination of the *VRRoom* and *VRTable* design spaces are preferable for specific tasks. One has to keep in mind that excessive interaction and switching between the two design spaces could impair some of the benefits, such as improved spatial memory capabilities. For 3D scatterplots, Yu et al. [425] presented a toolset of effective selection techniques in 3D pointclouds. In future research, such advanced techniques for the accurate selection of clusters could be implemented to assess if participants found the entire cluster. Also, advanced techniques that support the detection of clusters could be deployed, such as highlight-planes presented by Prouzeau et al. [295]. Besides the impact of interaction, it would be interesting to assess properties of the screen deployed. For instance, a larger screen with higher resolution might lead to higher levels of perceived immersion and increase task performance.

A larger sample size would have been beneficial to assess every experimental side factor accurately. However, we argue that the experimental side factors (noise, shape, density) were mainly deployed to guarantee the stability of results in the analysis of the visualization design spaces. We analyzed them as additional factors, but set the focus on the comparison of results for different visualization design spaces. Even though there was an exhaustive training session, and statistically no difference between experts and non-experts emerged, different outcomes could have emerged if we had conducted the study only with experts. We deployed two different kinds of datasets with regard to their noise level, much higher levels of noise could have changed the performance of users differently in each visualization design space.

One major limitation of our study is the restriction to three-dimensional data, favoring 3D design spaces, and thereby introducing a bias. However, we argue that our foundational research is targeting cases where dimension reduction to two dimensions is impossible or not advisable (e.g., see Figure 4.3). For truly high dimensional data a projection to 3D space might not make sense for cluster identification. We chose to focus on three dimensions as this is the maximum number of dimensions that can be encoded by the visual variable ‘position’ at a time in all deployed design spaces. However, this favors the 3D scatterplot visualizations as, for instance, if more than three dimensions had been represented in the visualization, a scatterplot matrix would have outperformed the three-dimensional scatterplots due to its dynamic scalability with regard to the number of dimensions. In addition to that, the 2D design space was disadvantaged as a higher learning curve can be expected for scatterplot matrices. More than half of the participants had none or few experience with scatterplot matrices.

Especially in the domain of molecular biology, 3-dimensional representations of molecular surfaces are often used, e.g., to investigate the size of genes, to compare proteins, or to identify substructures in electron tomography [195]. These spatial structures are comparable to point clouds visualized in 3D scatterplots. Therefore, we expect our results to be also true for similar tasks in such settings. Similarly, our findings could be applicable for applications with flow visualizations [379], spatio-temporal visualizations [5] and graph visualizations [179] in which entities have to be identified in a large 3D environment. Although not significant, participants performed better in the *VRTable* condition compared to *VRRoom*. When analyzing the subjective feedback, participants reported that they were missing an overview of the data when being entirely immersed in the *VRRoom* design space. We expect this circumstance to be independent of the visualization technique used. As a consequence, researchers should think about techniques to provide an overview of the data in VREs

To generally assess the possibilities of abstract visualizations for VA purposes in VR, future research should compare specific scenarios (task + visualization + data) in various visualization design spaces. The ultimate goal would be to establish some rules of thumb, advising one to avoid certain VA tasks and visualizations in VR design spaces and to favor the usage of others.

4.1.9 Conclusions

We presented a user study with 18 participants examining differences between four visualization design spaces with regard to cluster identification in scatterplots. The four employed spaces differed in their degree of immersion as confirmed by an additional study. Two of the design spaces were observed using a standard computer monitor (2D and 3D spaces on screen) and two using VR HMDs (restricted area in VRE and entire VRE). While the results show that more immersive visualization design spaces generally fit better to the given task, a fully embracing analysis environment may not be the best choice for scatterplot analysis due to a lack of overview and blind spots. Hence, for cluster identification tasks in scatterplots results suggest favoring a restricted area in a VRE as visualization design space. It is difficult to give a general recommendation when to use screen-based design spaces and when to deploy HMDs. We found that for scatterplot visualizations it can be beneficial to convey information by using three-dimensional VR design spaces if the task is to identify clusters in three-dimensional data. Results imply that thereby memory and orientation capabilities are increased. In comparison to abstract representations, 3D visualizations tend to be more comprehensive (maximally by using stereoscopic perception) and therefore ease the identification of clusters. However, abstract visualizations deliver more detail on single points or groups of points as extracting exact information from 3D visualizations can be difficult for humans due to distortion and a missing common baseline for comparing values that refer to multiple axes. Overall, we can state that VREs can indeed provide suitable design spaces for abstract visualizations such as scatterplots. Moreover, it became apparent that getting an overview of three-dimensional data can be enhanced by means of VR due to a more natural navigation, and better orientation and memorability capabilities.

4.2 A Comparative Study of Orientation Support Tools in Virtual Reality Environments with Virtual Teleportation

Movement-compensating interactions like teleportation are commonly deployed techniques in virtual reality environments. Although practical, they tend to cause disorientation while navigating. Previous studies show the effectiveness of orientation-supporting tools, such as trails, in reducing such disorientation and reveal different strengths and weaknesses of individual tools. However, to date, there is a lack of a systematic comparison of those tools when teleportation is used as a movement-compensating technique, in particular under consideration of different tasks. In this chapter, we compare the effects of three orientation-supporting tools, namely minimap, trail, and heatmap. We conducted a quantitative user study with 48 participants to investigate the accuracy and efficiency when executing four exploration and search tasks. As dependent variables, task performance, completion time, space coverage, amount of revisiting, retracing time, and memorability were measured. Overall, our results indicate that orientation-supporting tools improve task completion times and revisiting behavior. The trail and heatmap tools were particularly useful for speed-focused tasks, minimal revisiting, and space coverage. The minimap increased memorability and especially supported retracing tasks. These results suggest that VR systems should provide orientation aid tailored to the specific tasks of the users.

4.2.1 Introduction

Virtual reality is currently a popular trend in several areas: a developing private market in the computer games sector, in the industry for business solutions, but also in research – for example, in the fields of human-computer interaction and data visualization. While the virtual environment can be of any size, the ‘play area’ – i.e., the physical space available in the real world – is usually much smaller, leading to a discrepancy between the two worlds. Naturally, one would walk around for spatial navigation, leading to improved orientation abilities due to self-motion [191, 313]. However, often the available physical space is limited, confining this interaction in virtual reality. Razzaque et al. [303] presented ‘redirected walking’ as a technique for enlarging the virtually accessible space by tricking the orientation and perception of users. In their approach, the virtual environment is rotated, unnoticed by the user, detaching the spatial mapping of the real world from the virtual world. Thus, the user’s physical movement can be limited to a small area, while the virtual range is (theoretically) unlimited. However, to keep physical direction changes unnoticeable to the user, it is necessary that the user walks in wide arches, which makes the technique unsuitable in confined physical space (e.g., 3×3 m).

Another way to alleviate the problem of limited physical space is to intercept the physical input (i.e., walking) and simulate it in the virtual environment. So-called VR treadmills [277] are large installations, on which a VR user can walk on the spot, covering virtual space while staying in the same physical location and still using natural interaction as a means of navigation. However, the approach has disadvantages: Besides a bulky setup and expensive hardware, the approaches are not yet fully developed and do not perfectly mimic real interactions [350].

Indirect spatial navigation via buttons or joysticks, typically used in computer games, is an alternative way of overcoming the differences between the virtual and the real world. However, this type of movement fosters an asynchronous movement between body and perspective, which is unfamiliar to our brain and can, thus, lead to physical discomfort – often referred to as simulator or motion sickness [154, 217].

Another alternative is virtual 'teleportation', which circumvents the aforementioned problems associated with asynchronous movement. The user points to a specific location within sight and triggers the movement with a controller or similar. The user's location is then abruptly set to the selected position. Initial research suggests that this might not be the ultimate solution either, as it fosters a loss of orientation, since the point of view is often changed without physical transition of the user [30, 410]. Various orientation aids can be deployed to counteract such losses of orientation, such as trails or minimaps.

In this chapter, we seek to better understand how well different visual orientation-supporting tools compensate this disorientation problem by evaluating and comparing three typically used tools collected from literature. *Orientation-supporting tools* are means helping us to maintain our orientation in (VR) environments, even when it is made difficult by unfamiliar movement techniques such as virtual teleportation. Navigational tools can be considered as a subgroup of orientation-supporting tools. Navigational tools require the destination to be known so that the system can help the user to find a way to this target location. We focus on environments where the physical space is limited, and teleportation is deployed as a movement-compensating method. We selected three techniques for evaluation. Two of them are frequently used for facilitating orientation: a minimap and trailblazing. As a third technique, we investigated the performance of real-time heatmaps to support orientation. In a user study with 48 participants, we compared these three approaches to each other and to a baseline scenario without any orientation-supporting tools.

The main contributions of our research are *(i)* the results of a between-subjects user study that assesses and compares three passive and stand-alone orientation-supporting tools in VREs with limited physical space that use teleportation for spatial movement, and *(ii)* a set of guidelines that propose types of tools for different scenarios and tasks.

4.2.2 Background

In the following, we provide definitions of prominently used terms, as well as an overview of various application scenarios in which navigation and orientation-supporting tools play a role. Additionally, we give an overview of existing techniques and categorize them according to distinguishing properties.

4.2.2.1 (Digital) Wayfinding and Orientation

Darken and Peterson [82] define *navigation* as a combination of *wayfinding* and *motion*, with *wayfinding* being the cognitive element and *motion* the motoric element. Further, *spatial cognition* is "concerned with the acquisition, organization, utilization and revision of knowledge about spatial environments" [123]. Spatial *orientation* is associated with spatial cognition and describes the ability to orient oneself relative to specified positions on a cognitive map. A higher degree of orientation increases the performance in wayfinding and navigation tasks [123, 233]. Orientation is a complex construct that can affect many aspects that rely on one's abilities to grasp and understand the spatial structure of an environment and its spatial relationships. In the current evaluation, we make use of three indicators for improved orientation: the ability to gain and maintain an overview of the environment, the ability to return to previously visited places, and the ability to maintain an overview of all previously visited places.

While it is apparent that orientation is important in navigation scenarios where the main task is to find an (optimal) path from one location to another, it is also important in other scenarios without concrete navigational tasks. For example, in visual exploration scenarios in which the user navigates through a large data space, it is important that analysts can maintain their orientation. Examples are the exploration of multidimensional data visualized as scatterplots in immersive environments [96] or the exploration of digital models of cities [432], buildings [292] or crime scenes [402].

Previous work has already investigated in detail how navigation can be supported by providing additional visual aid. Thereby, the research is based on real-world applications where technology was developed and evaluated to guide users through real environments. A classic example is the navigation system in a car. The location of the car is shown on a map and a route is highlighted that suggests where the driver should go. Besides, there is a large number of research projects dealing with navigational support in digital environments; i.e., screen-based applications where users are guided from one location to another [83, 87, 318]. The prevalent examples are gaming and simulation applications where a user navigates in a virtual real-world simulation.

4.2.2.2 Improving Orientation in VR

Another branch of research deals with *orientation* and how it can be improved in virtual reality environments. In this area the ultimate goal is to develop techniques that support users in improving their cognitive map of an environment.

Virtual reality environments are frequently used to study basic human navigation and orientation capabilities in laboratory setups [237, 253, 378]. For instance, Moffat et al. [253] examined the influence of age on participants' ability to navigate in an abstract labyrinthine environment and found an impairment of orientation that correlated with age. Similarly, Maguire et al. [237] investigated the influence of age and gender on orientation capabilities and found indications for interactions between the two properties and orientation, for example, lower orientation skills with increasing age.

Virtual reality is a relatively new trend that is constantly expanding into new applications. Although the environments explored are digital, and in this respect similar to screen-based digital environments, due to immersion, navigation in VR is more similar to navigation in the real world. Therefore, various techniques for navigation and orientation support were transferred from screen-based applications to VR and partially evaluated, for example, trails, compasses, and minimaps [3, 83, 84].

Apart from that, techniques were developed and evaluated that are tailored to stereoscopic vision and 3D space provided in virtual reality environments, such as 3D minimaps [45], 3D radars [44], edge radars [144], or fisheye views [326]. Kotlarek et al. [194] compared landmarks, 3D minimaps and waypoint navigation with regard to their ability of improving spatial orientation and identified the 3D minimap as the most efficient aid in their comparison. Besides providing explicit visual tools for orientation improvements, several studies considered more indirect ways to improve orientation. For example, Müller et al. [257] describe how collaborative VR environments can be configured in order to foster orientation. They found that shared virtual artifacts increased orientation in remotely co-located virtual environments for collaborative tasks.

Riecke and Schulte-Pelkum [313] identified self-motion as a crucial element for orientation and set out how the illusion of self-motion (vection) could be used in virtual environments to enhance orientation. This is in particular important for VREs with limited physical space as they commonly rely on alternative locomotion techniques to walking. The impact of using movement compensating techniques, such as teleportation [30] or controller-based movement [410] on orientation has been investigated in previous studies. Bhandari et al. [23] presented 'Dash', an alternative technique to teleportation in which a user is slowly transitioned to a selected location and compared it to its original alternative. They found that provided optical flow cues in their new technique increased orientation. Similarly, Langbehn et al. [217] compared redirected walking, joystick navigation and teleportation and found an advantage of redirected walking compared to the other two approaches, as it helps the user to unconsciously acquire spatial knowledge of the virtual environment.

However, in many VR applications the use of standard virtual teleportation still predominates. Therefore, we complement to the line of research of finding suitable tools to counter orientation loss induced by virtual teleportation by comparing three well-established, visual orientation-supporting tools in virtual reality environments.

4.2.2.3 Types of Visual Orientation Support

We categorize existing techniques of visual orientation-supporting tools according to three characteristics: environment dependency, target dependency, and action dependency.

Environment dependency: Environment-dependent tools are tools where the environment has to be attuned to the respective task. An example of an environment-dependent orientation-supporting technique is *sectioning* [86, 324]. In this approach, the environment is divided into visually distinct segments. For example, in the virtual model of a building each level could have a carpet in a different color. The approach of *landmarking* [324, 355, 417], in which visually distinguishable landmarks (e.g., objects as salient cues) are placed in the virtual environment to serve as orientation anchors, is also environment-dependent. With both approaches, the environment has to be adapted beforehand. Environment-independent or *stand-alone* approaches, on the other hand, are applicable without manipulating the environment itself. Examples are minimaps or static directional cues, such as a pointer to the north or a virtual sun [82].

Target dependency: Any navigational support that guides the user to a system known location is target-dependent. Conventional navigation guides, such as minimaps with a given target location or even a suggested path, are examples of this category [82], as well as radars or arrows indicating target locations [44, 63]. Target-independent approaches are all approaches that do not take the target location into account, such as the *breadcrumb* or *footprint* technique, which depicts a person's travel history on the ground [138, 318].

Action dependency: Action-dependent approaches require *active* user interaction. The predominant example is interactive landmarking [66]. With this approach, the user can place different objects as spatial cues in the virtual environment to improve his or her orientation. Any other *passive* tool that does not require active user interaction is categorized as action-independent.

In our work, we focus on environment-independent, target-independent, and action-independent orientation-supporting tools in virtual reality environments, as they are the most versatile aids for exploration tasks.

4.2.3 Assessing Orientation-Supporting Tools in VR

In this chapter, we assess and compare three different orientation-aiding tools in VREs. We focus on VREs with limited physical space, which therefore use teleportation as a means of spatial movement. To test our assumptions, we conducted a quantitative user study. The study prototype was developed with the gaming engine Unity3D [371]. In the following, we will first present the rationale behind the conducted evaluation and justify the selection of three orientation-supporting tools before continuing with a detailed description of the conducted user study.

4.2.3.1 Research Objectives & Study Design

Teleportation is a technique frequently used in VREs to overcome limited physical space while avoiding the deployment of expensive hardware or hazarding the risk of increased levels of simulator sickness caused by asynchronous body and viewpoint movements. Bowman et al. [30] found that the use of teleportation for spatial movement in virtual reality environments has an impact on spatial orientation and suggest the deployment of alternative ways of relocation instead of teleportation. Our goal is to investigate possible solutions for the loss of orientation while maintaining the popular technique of teleportation for movement. There are several techniques that are developed to support a user's orientation. Our key research questions are: *Do common orientation-supporting techniques indeed help users with orientation? Which technique is best suited to maintain a user's orientation in a complex environment that promotes disorientation?*

The conducted study comprised only one experimental factor: the provision of a specific orientation-supporting tool. Since our goal was to compare three different techniques with each other and with the provision of no tool at all, trials were carried out in four conditions. The study design was a between-subjects design and included 48 participants in total. Thus, the sample was divided into four groups of 12 participants, each group using one of the three orientation-supporting tools or none at all.

We used mazes of different sizes as abstract environments to promote disorientation. However, to test the applicability of our findings, we also used a session in which participants performed a search task in a realistic environment. To investigate the orientation of users, we used different exploration and search tasks that allowed us to quantify orientation.

We focused on the comparison of the orientation-supporting tools in their basic form, without considering any combinations of them. Based on these initial results, future research should evaluate the pairings of approaches as well. As far as the selection of examined techniques is concerned, we selected well-functioning and optimized techniques from other domains or evaluations that meet our requirements.

4.2.3.2 Conditions: Orientation-Supporting Tools

We evaluated four orientation-supporting tools in our study.

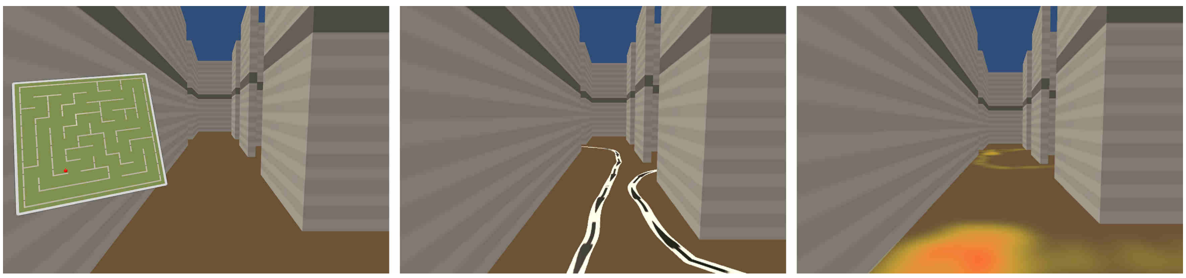

Figure 4.10: Provided orientation-supporting tools. Minimap (left): Attached to the left controller, a north-up minimap provides the user with an overview. The position and orientation of the user is depicted by a small icon on the map. Trail (center): As the user moves through the space, a trail is created highlighting the path taken. Heatmap (right): A heatmap on the floor depicts the history of user locations.

None In this baseline condition, participants completed the same trials as the other participants but were not equipped with any orientation-supporting tool. This condition was introduced to check whether the supportive tools have any positive benefit for user orientation and to serve as a baseline for comparison.

Minimap *Minimaps* are frequently used navigational and orientation-supporting aids. In the case of our study, we wanted to focus on orientation-supporting aids. Hence we only used scenarios where the target location was unknown. Following the design principles of Darken and Peterson [82], we used a north-up map, which is optimized for geocentric tasks like explorations or naïve searches (see Figure 4.10, left). For egocentric tasks, such as navigating to a known location, forward-up maps performed better in previous evaluations [84]. To ensure optimal performance of the map, we added a user icon that always indicates the user's position and orientation on the map, as suggested by Darken and Peterson [82]. In the study, the minimap was attached to the user's left controller and could be moved around as desired. An orthographic camera was centered above the respective scenario, so that the user-accessible bounds were aligned with the bounds of the map. This means that the entire environment was continuously visible on the minimap. The result can be compared with a map of the ground plan of the scenario. Task-based markers were not depicted in the minimap.

Trail As shown in Figure 4.10 (center), the orientation-supporting tool *Trail* draws a trail behind the user while he or she moves in the virtual space. Arrows on the trail indicate the direction in which the path was taken. Based on Darken and Sibert [86], we decided against a breadcrumb approach and instead used a continuous line with directional cues – a directed trail. Continuity makes the trail more robust against overlaps and crossings, as opposed to, for example, footprints [138]. Directional cues, such as textures on the trail that indicate the direction in which the path was taken, increase the user's ability to keep track of the path, and reconstruct the history of movements. Darken and Peterson [82] express concern about the use of trails as they clutter the space after long sessions, but at the same time point out that the alternative of fading trails can lead to confusion because the aid no longer provides binary information about whether or not a particular place has been visited before. As we designed our sessions to be short (about 5 minutes) and used environments with a small areas (maximum of 100 m^2), we retained all trails from the beginning to the end of a session. Furthermore, we abstained from visually encoding the recency of a trail, for example, by color, as suggested by Ruddle et al. [318], since this would lead to non-persistent visualizations that change color over time, which in turn could lead to a loss of orientation. Regarding the technical implementation, we used a textured line renderer, that is extended by waypoints as the user moves. We set a threshold of 20 cm to the latest point so that no new waypoints are created unless the user moves significantly. To smooth the corners of the displayed line, we implemented an interpolation function that makes the displayed trails more visually appealing. Normals of the lines are set to the normals of the ground plane. Finally, we set a threshold of waypoints for the line renderer. Once the threshold is reached, the line is converted to a mesh, and the line renderer is reset to improve performance.

Heatmap The fourth scenario was inspired by applications that use heatmaps in post-processing steps to analyze user behavior. Heatmaps were used, for example, in gaze analysis procedures to visualize the eye movements of users [117, 358], and in sport analysis scenarios to analyze the movement of players in a soccer game [17, 281]. We attempted to transfer the benefits found for such applications by using them as real-time updating heatmaps for Orientation support: Unlike trails, heatmaps do not encode the direction from which a place was visited, but only the frequency with which that place was visited. This reduces the complexity of the presented visualization at the price of information loss. To implement the technique, the total area was divided into 10×10 cm tiles on a $n \times m$ grid. Each tile served as a counter and monitored how long the player was standing on it. Every 0.5 seconds, the counter of the tile beneath the current player position was increased by one. To generate the heatmap, a $n \times m$ matrix was created, consisting of all counters. We then linearly normalized all non-zero values to a range between 0.1 and 1. This normalization prevents spots in the heatmap from disappearing due to very high counts at certain positions. As a next step, we created a $n \times m$ texture from the matrix by assigning a color to each value based on our color gradient. As a color gradient, we deployed a linear color transition from fully transparent to yellow to red. Finally, we applied a Gaussian blur function [147] to the texture for smoothing. This Gaussian filter is a $N \times N$-tap convolution filter that weights all pixels in its catchment area based on the Gaussian distribution function. To visualize the heatmap, a texture plane was placed 1 mm above the floor. The texture of this plane was updated every two seconds (see Figure 4.10, right).

4.2.3.3 Hypotheses

Based on observations from a pilot study ($N = 4$, same study prototype) and findings from relevant literature, we derived the following hypotheses. The evaluated scenarios were tailored to our hypotheses, focusing on a specific task type (exploration, naïve search, and informed search) and related task objectives (completeness of exploration, overview, and retrace abilities).

H1 We expect improved orientation when orientation-supporting tools are available. In general, we expect that participants are more efficient and more effective in exploration, naïve search and informed search tasks with all given tools, as previous research has shown a benefit of such tools in screen-based navigation applications [86]. This should be reflected, among other things, in lower task completion times, higher task accuracy for exploration tasks, and lower navigation times for informed search tasks.

H2 We expect that the orientation-supporting tool *Heatmap* increases the effectiveness in open exploration tasks. The heatmap indicates previously visited areas and thus implicitly points to areas that were not visited before. In contrast to the *Trail* visualization, it does not create clutter but provides a higher abstraction of the movement history. This should be reflected in a higher task accuracy for exploration tasks. Participants with a *Heatmap* are expected to leave less space unexplored and thus find more targets.

H3 We expect that the orientation-supporting tool *Minimap* offers the best overview in naïve search tasks. The *Minimap* provides a direct top-down view of the environment, revealing the structure of the environment and the user's position [86], helping the user to quickly gain and maintain an overview of the scenario to be explored. Therefore, we expect participants with a *Minimap* to be able to find a certain set of targets faster than participants without or with another tool when performing a naïve search task. In addition, we expect that participants are able to recall the structure of the environment better if they are provided with a *Minimap*.

H4 We expect that the orientation-supporting tool *Trail* increases tracing performance in informed search tasks - i.e., the ability of a user to navigate back to a known previously visited location. Ruddle et al. [318] demonstrated that this technique can be effective for subsequent search tasks. We, therefore, expect participants with the orientation-supporting tool *Trail* to find their way back to known locations fastest.

4.2.3.4 Procedure

Participants were invited to the laboratory for individual sessions. Prior to participation, all participants gave written informed consent. The experiment was divided into three parts. In the first part, participants filled in two standardized questionnaires assessing their mental rotation abilities and basic orientation skills: the Mental Rotation Test (MRT) by Peters et al. [282] and the Perspective Taking and Spatial Orientation test (PTSO) by Hegarty and Waller [151]. Such spatial ability tests are recommended by Darken and Peterson [82] because they make it possible to relate analysis results to participants' spatial abilities. In the second part (main experiment), participants first completed a training procedure in VR in which all available interaction techniques were introduced (how to open markers, how to teleport). Subsequently, they performed four trials and answered questionnaires after each trial. Before each trial, participants took a short break and completed a calibration and instruction step in which they positioned themselves in the starting position (center of physical room) and read the instructions for the following trial. We provide a video, in which the four trials, as well as all conditions are graphically presented[1]. After the four trials were completed, the third part began, consisting of a customized questionnaire, a demographic questionnaire, and a semi-structured interview. Finally, participants were thanked and compensated for participating.

4.2.3.5 Interaction Capabilities

Regardless of the trial, participants could walk in the physical space available (2.5×2.5 m). Additionally, participants could press and hold a button on the controller to select a location to which they were teleported after releasing the button. As a second interaction, participants could interact with markers by pointing at them and pressing a button on the controller. This interaction was only possible from a distance of 0.5 m. Once a marker was triggered ('opened'), it changed its appearance.

4.2.3.6 Tasks & Trials

Darken and Sibert [86] distinguish between three types of tasks: exploration, naïve search, and informed search. We focused on the first two types, as we strove to investigate the performance of orientation-supporting tasks where the target location is not known to the system or the user. However, in order to also investigate user performance in informed search tasks, we constructed a naïve search task in which the user has to navigate back (retrace) to previously found targets, resulting in an informed search sub-task. Throughout the study, participants were asked to collect markers in four trials. However, the main objective of the specific task varied from trial to trial, defining its task type. The order of the trials was the same for all participants as we did not conduct cross-comparisons between tasks and strove to

[1] `https://youtu.be/rnPvBTJQeh0`

maintain the same learning effect for each participant in each task. In the following, we will elaborate on the descriptions of each trial.

Trial 1: Maze Exploration The main objective of exploration tasks is to explore as much of the environment as possible without leaving areas unexplored. Trial 1 can be seen as an exploration task, as participants were asked to explore the respective scenario without knowing how many markers there are and when they are finished. Participants were instructed to explore the entire scenario carefully, as thoroughly as possible, and to collect all available markers.
In the first trial participants were located in a 15×15 m maze (see Figure 4.11, top left). Four markers were distributed in the maze. Participants were instructed to explore the maze and visit every corner of the maze. Moreover, they were instructed to collect all markers in the maze and to return to the starting position (blue circle on the floor) as soon as they are confident they have explored every inch of the maze. After this trial, participants filled in the NASA Task Load Index (TLX) questionnaire [148].

Trial 2: Naïve Search Task The main objective in naïve search tasks is to find a set of targets efficiently. Trial 2 can be seen as a naïve search task as participants were asked to search for a certain number of markers. Participants were instructed to collect all four markers as fast as possible and return to the starting position.
The second trial took place in the same maze as the first trial. However, the four markers were placed in other locations. Participants were made aware of both circumstances. We used the same maze to exploit learning effects and to measure whether certain orientation-supporting tools enhance participants' mental map of the environment and thus facilitate the task. After this trial, participants again completed the TLX and a memorization questionnaire, assessing participants' ability to remember the structure of the maze.

Trial 3: Informed Search Task The main objective in informed search tasks is to efficiently navigate to known locations. In trial 3, participants were asked to collect markers in a specific order and return to the starting position. As soon as a marker was 'opened', it revealed a certain number (1, 2, or 3). We manipulated the numbers on the markers so that participants always found the third marker first, then the second, and then the first marker. So when they finally got to the first maker to collect, they knew the position of the other two markers. In this way, we were able to construct an informed search task that starts with a naïve search task where neither the system nor the user knows the targets in advance.
In the third trial, participants were located in a smaller, 10.5×10.5 m maze (see Figure 4.11, bottom left). Three markers with numbers (1, 2, and 3) were distributed in the maze. After this trial, participants filled in the TLX, a memorization assessment questionnaire, and the simulator sickness questionnaire (SSQ) by Kennedy et al. [184].

Trial 4: Real-World Inspired Exploration In the last trial, we used a more realistic scenario (see Figure 4.12) in which participants performed a similar exploration task as in trial 1. In the approximately 1,000 m^2 large walkable space, ten 1 cm large diamonds were distributed. Participants were instructed to explore the environment for five minutes and collect all diamonds. After this last trial, participants again completed the TLX questionnaire.

4.2.3.7 Scenarios / Virtual Environments

For our study, we created two types of VREs: abstract and realistic.

Abstract Scenarios: Mazes of Trials 1-3 Darken and Sibert [86] found that people automatically take advantage of environmental cues and partition spaces for better orientation. We decided to minimize this benefit of natural-looking environments by designing the environments for the first three trials as monotonous and uniform as possible. In addition, we used mazes as they naturally promote disorientation, thus increasing the potential impact of orientation-supporting tools.

Figure 4.11 depicts the ground plans of both mazes used (left) and a screenshot of the first-person perspective of a participant in the maze (right). The blue dots in the ground plans represent participants' starting positions. For the first two trials, a larger maze with a dimension of 15×15 m was used. For the third trial, the dimension of the maze was 10.5×10.5 m. We used the same mazes for all participants to be able to compare their performance. The maze was not changed from the first to the second trial in order to exploit learning effects and to familiarize participants with the maze in the exploration task (trial 1) before they continued with the naïve search task of trial 2.

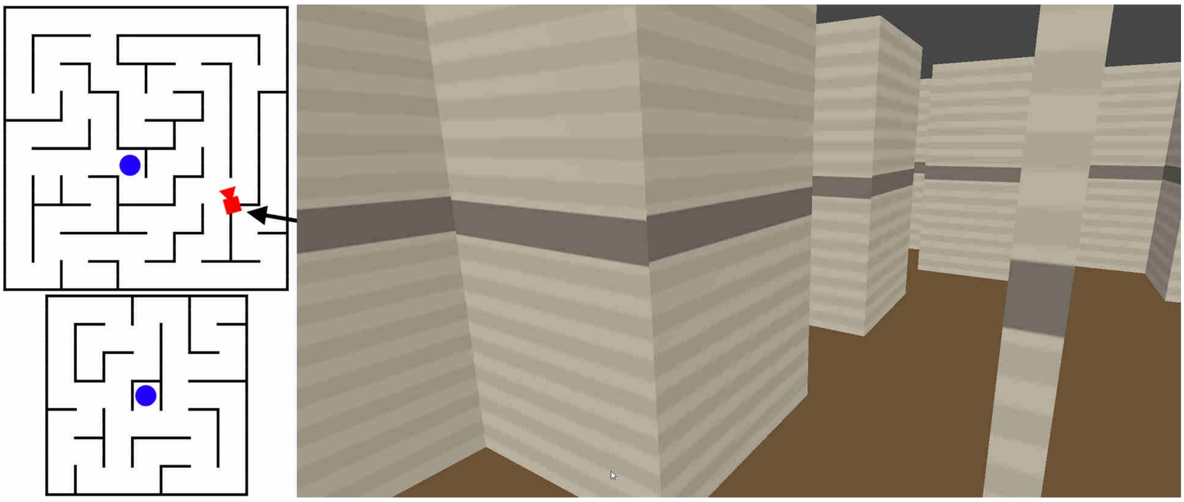

Figure 4.11: Mazes used in trials 1-3. Left: ground plans of mazes for trial 1 and 2 (top) and trial 3 (bottom). Right: sample image from the perspective of a participant standing in a maze.

The two mazes used were created randomly using a depth-first search algorithm. First, the ground plan of the maze was divided into 1.5×1.5 m cells, and each cell was framed with four walls. One cell was randomly selected and marked as visited. For each cell currently visited, a list of neighbors was retrieved and processed in random order. For each neighbor not yet visited, the walls between the two were removed, and the same function was recursively called for this cell. We then removed various walls manually to retrieve unconnected walls. The sizes of the mazes used were obtained from pilot studies in which we measured the time participants needed to complete the given tasks in the respective mazes. We chose the mazes in which the average participant required 5 minutes to solve the task. Since the task in the third trial was more extensive compared to the first two trials, we decided to use a smaller maze for this trial. The texture of each maze was uniformly striped to enhance depth perception and the local spatial structure recognition of the maze, as participants in the pilot studies had difficulties in recognizing the corners of a monochrome maze.

Realistic Scenario: Industrial Environment of Trial 4 Since mazes are rather unusual environments, we extended our study to include a more realistic use case that is inspired by a real-world scenario. Therefore, we created a fourth trial in which participants should explore a more realistic environment. For this, we used a 3D map of a fictional industrial environment. As a base environment, we adopted a map by Dimitrii Kutsenko [212] from the Unity3D asset store and adapted it slightly to fit our needs. Figure 4.12 depicts a top view of the scene (left) as well as a screenshot from the first-person perspective of a participant in VR (right). The red area marked in the left image illustrates the area that was accessible for the participant. The blue spot in the right image represents the starting position of the scene. The approximately 1,400 m^2 comprise obstacles, such as pallet stacks, barrels, silos, and even houses, resulting in a total accessible area of about 1,000 m^2.

Figure 4.12: Scenario used for the real-world inspired exploration task (trial 4). Left: top view of the map. The red boundary indicates the area in which participants could move. Blue dots indicate participants' starting position. Right: one image from the participants' perspective. The location from where this image was taken is marked as a small red camera in the left image.

Markers Trials 1 to 3 were designed with large markers (Figure 4.13, left). The marker consists of a 35 cm cube with a 25 cm long diamond on top. Before interacting with the marker, the box was black with a red question mark, and the diamond was grey. Once opened by a user, the appearance changed. As the order of markers was irrelevant for trials 1 and 2, the diamond turned green and the box blue (see Figure 4.13, left). For the third trial, the order in which the markers had to be opened was relevant. Therefore, the box showed the identifier of the opened marker, and the diamond indicated whether the opened marker had to be opened next (green texture) or not (red texture), as depicted in Figure 4.13 (center). If the marker was not the right one, it returned to its original appearance after five seconds.

In the fourth trial, we used different and much smaller markers as the large ones were far too easy to recognize in the plain environment (see Figure 4.13, right). The markers used were 9 mm long diamonds of grey color. When opened, the markers were enlarged to a length of 5 cm and changed their color to green. Before the start of the trial, participants were made familiar with the appearance of the diamonds to be sought.

4.2.3.8 Apparatus

The experiment took place in a laboratory at the University of Konstanz. In addition to the participant, a study supervisor was present at each session. The physical space in which participants could move

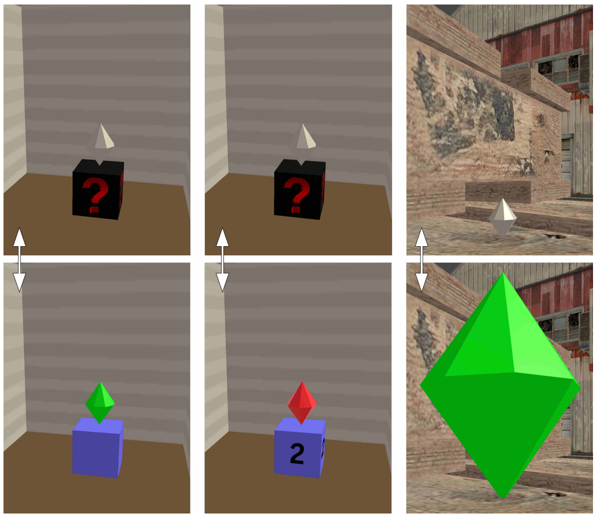

Figure 4.13: Used markers in the search and exploration tasks. Left: large markers of trials 1 and 2. Center: large markers with numbers of trial 3. Right: tiny markers of trial 4 (9 mm before opening, 5 cm after opening). The top image of each pair of images depicts the marker before opening, and the bottom one represents the marker after opened by a participant.

was an area of 2.5×2.5 m. Participants were equipped with an HTC Vive Pro [161] and two Vive controllers. To ensure maximum mobility, we used a wireless adapter for the Vive HMD.

4.2.3.9 Sample

A sample of $N = 48$ participants (33 female, 15 male; aged between 18 and 43 years) was recruited using the study pool of the University of Konstanz. Participants were sequentially assigned to one of the conditions depending on their gender to counteract possible gender differences in spatial orientation [70]. Median ages per condition were as follows: None ($Mdn = 23.5$), Minimap ($Mdn = 24$), Trail ($Mdn = 25$), Heatmap ($Mdn = 23$). Most participants did not have much experience with VR ($M = 1.40$, $SD = 0.22$) and computer gaming ($M = 2.15$, $SD = 0.32$). Means represent experience ratings of users on a scale from *very few* = 1 to *very much* = 5.

4.2.3.10 Results

We report significant results of our quantitative analysis as well as qualitative feedback from the final interview and the customized questionnaire. Statistical analysis was performed using IBM SPSS Statistics (version 25). All tests are based on a 0.05 significance level. Each dependent variable was first tested for normal distribution by the Kolmogorov-Smirnov test. We computed a one-way independent ANOVA for normally distributed data in combination with a Tukey-HSD test as a post-hoc test. For non-normally distributed data, we used its non-parametric counterpart, the Kruskal-Wallis test, and a Mann-Whitney test as post-hoc test. Graphs display mean values, with error bars indicating standard errors of the mean. Asterisks indicate significant differences (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

Before the trials started, participants' basic mental rotation and orientation abilities were assessed with two standardized tests (MRT [282] and PTSO [151]). Statistical analysis did not reveal significant differences between groups, i.e., participants in the four conditions did not differ in their mental rotation or orientation skills (MRT: $H(3) = 2.4, p = .494$; PTSO: $F(3,44) = 1.49$, $p = .230$).

Task Completion Time Task completion time was measured as the time difference between the start of the trial and the time of return to the starting position after completing the task. Since the time in trial 4 was limited to five minutes, we consider the task completion times only for trials one to three. For the maze exploration (trial 1), we corrected task completion times to the interval from the start of the trial to the last marker found as it was an open exploration task without a designated number of targets. Data for all three tasks considered were normally distributed.

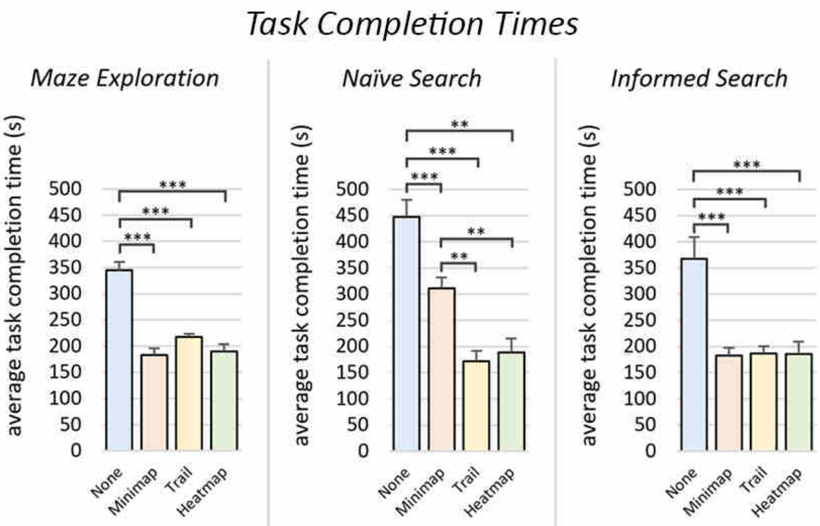

Figure 4.14: Average task completion times, i.e., the duration from the start to the end of a trial.

As depicted in Figure 4.14, task completion times differed significantly between groups for all three tasks: *Maze Exploration* ($F(3,44) = 31.9$, $p < .001$, $\eta^2 = .68$), *Naïve Search* ($F(3,44) = 23.5$, $p < .001$, $\eta^2 = .62$) and *Informed Search* ($F(3,44) = 11.5$, $p < .001$, $\eta^2 = .44$). For all three tasks, Tukey-HSD post-hoc tests revealed that the condition without any orientation-supporting tool (*None*) was significantly different from all others ($p < .05$). Apart from that, only for the *Naïve Search* task completion times were significantly longer in the *Minimap* condition than both in the *Heatmap* and *Trail* condition.

Task Accuracy Task accuracy was measured as the percentage of markers found in the exploration tasks (trials 1 and 4). We only assessed the accuracy for these two tasks because the number of markers in the other two tasks was predetermined, and the task could, therefore, only be completed with all markers found. In the maze exploration, four markers were distributed in the maze, whereas in the 'real-world' inspired exploration, ten markers (diamonds) were placed in the more realistic environment. In the first task, data were normally distributed and statistical tests indicate significant differences ($F(3,44) = 5.3$, $p = .003$, $\eta^2 = .27$). For the fourth task, the non-parametric Kruskal-Wallis test reveals significant differences between used orientation-supporting aids ($H(3) = 15.30, p = .002, \varepsilon^2 = .33$).

As shown in Figure 4.15, in the maze exploration task participants found significantly more markers in the *Heatmap* and *Trail* condition than in the *None* condition ($p < .05$). The same applies to the real-world inspired exploration task, where in addition task accuracy was significantly lower in the *Minimap* condition compared to the *Heatmap* and *Trail* condition.

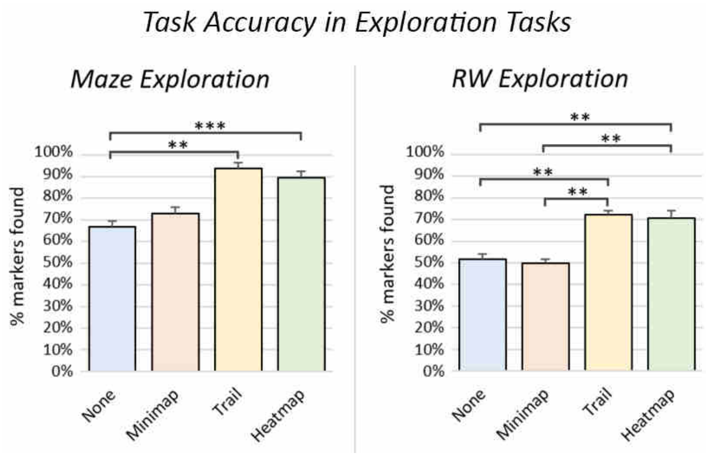

Figure 4.15: Mean values of task accuracy, measured as the percentage of markers found in the two exploration tasks (trial 1 and trial 4).

Spatial Exploration We assessed the spatial exploration of participants in two ways. On the one hand, we evaluated how much of the accessible space participants actually covered in the exploration tasks (trials 1 and 4), and on the other hand, how much space was redundantly visited in all tasks.

Explored area The percentage of space explored was calculated on the basis of participants' movement data. All accessible areas were divided into 1.5×1.5 m tiles. Once a participant visited a tile, it was marked respectively. In the maze exploration task (trial 1) the 15×15 m maze was divided into 100 tiles, whereas in the real-world inspired exploration task, the virtual compound of 1000 m^2 was divided into 1000 tiles. Data in the maze exploration task were normally distributed and indicate significant differences ($F(3,44) = 9.45$, $p < .001$, $\eta^2 = .39$). As shown in Figure 4.16, in the *None* condition participants explored significantly less space of the maze compared to participants in the *Heatmap* or *Trail* condition ($p < .01$). Additionally, there was a significant difference between the *Trail* and *Minimap* conditions. In the real-world inspired exploration task, data were non-normally distributed, and no significant differences between provided tools could be found ($H(3) = 2.32, p = .509$).

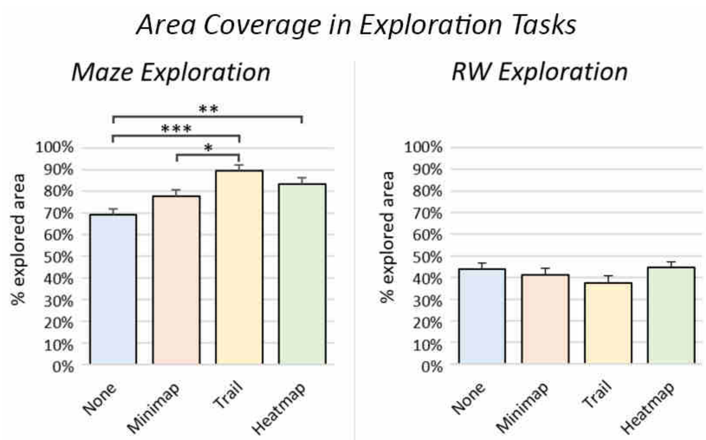

Figure 4.16: Mean values of exploration coverage of the accessible space in exploration tasks (trial 1 + 4).

Redundantly visited space To calculate the area that participants visited several times, we counted the tiles that registered the player more than once. Data were normally distributed only for the naïve search task and the real-world inspired exploration task. Statistical tests indicate significant differences for the naïve search task and the informed search task: *Maze Exploration* ($H(3) = 3.08, p = .380$), *Naïve Search* ($H(3) = 16.97, p = .001, \varepsilon^2 = .36$), *Informed Search* ($F(3,44) = 6.16$, $p = .001$, $\eta^2 = .30$), *Real-World Inspired Exploration* ($F(3,44) = .917$, $p = .440$).

As shown in Figure 4.17, in the naïve search task participants visited more areas redundantly in the *None* condition compared to all other conditions. In the informed search task this difference was only found between the *None* and *Heatmap* condition. However, in this task participants visited more space redundantly in the *Minimap* condition than in the *Heatmap* or *Trail* condition.

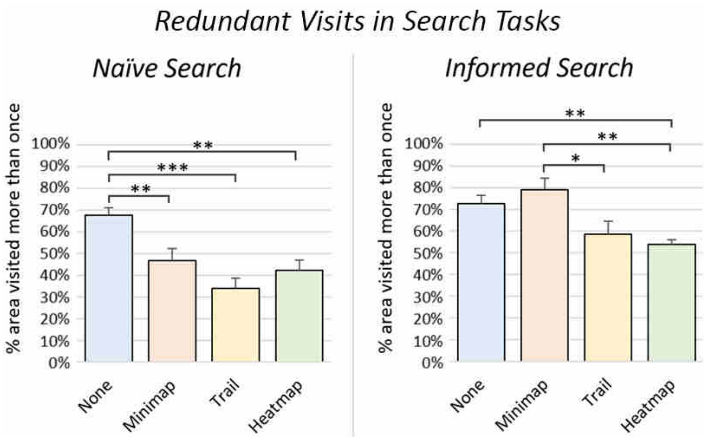

Figure 4.17: Mean percentage of areas that were visited more than once. Shown are tasks with significant differences: the naïve search task (trial 2) and the informed search task (trial 3).

Retrace Performance Participants' retrace skills were assessed in two ways. Firstly, in the naïve search and informed search tasks, as the time needed to return to the starting position after the task was completed. Secondly, only in the informed search task, as the time, distance, and error between two known markers. In this task, participants were asked to collect three markers in the right order. As participants always found the markers in the order 3, 2, and then 1, they 'knew' the location of the second and third marker as soon as they had collected the first marker. Therefore, we could measure the variables as mentioned above for these portions. We assessed the time it took from one marker to the next, the distance participants covered between them, and the error, i.e., how often they tried to collect the third marker immediately after they had collected the first one.

Analyzing the time, that participants needed to return to the starting position after the last marker was found, showed no significant differences for the informed search task ($F(3,44) = 1.76$, $p = .170$), but for the naïve search task ($F(3,44) = 3.6$, $p = .022$, $\eta^2 = .11$). In this task, participants needed significantly more time without an orientation-supporting tool (*None*) as if they were provided with the minimap (*Minimap*).

Considering only the informed search task, in which we assessed participants' ability to trace back their way to a previously found marker, for both retracings significant differences in time emerged: from the first to the second marker ($F(3,44) = 3.1$, $p = .037$, $\eta^2 = .17$) and from the second to the third marker ($F(3,44) = 4.1$, $p = .012$, $\eta^2 = .19$). Figure 4.18 depicts mean values of times participants needed to get from one marker to the next in the right order. In both cases the times in the *None* condition were significantly longer than in the *Minimap* condition ($p < .05$). Furthermore, the time for tracing from the second to the third marker was longer in the *None* condition compared to the *Trail* condition. Similar results emerged when considering the variable distance. The analysis of the error rates did not reveal significant differences ($F(3,44) = 0.55$, $p = .649$).

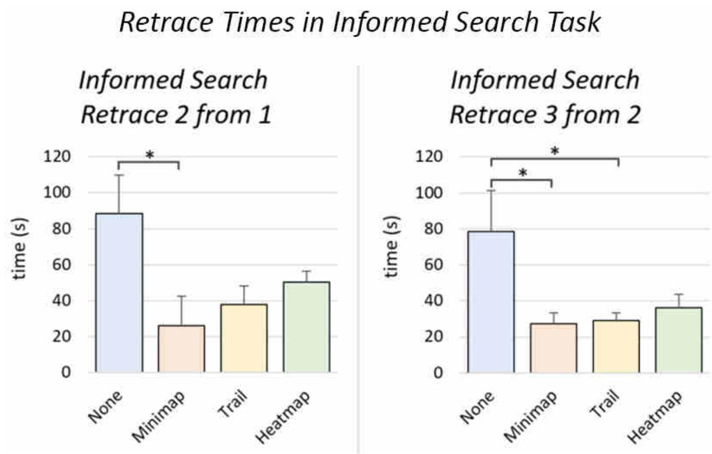

Figure 4.18: Mean values of retrace times, measured as the time it took participants to navigate from one marker to another in the informed search task (trial 3).

Memorization After trial 2 and trial 3, participants were asked to choose the correct ground plan of the maze they previously had walked through from three options. Since the first two trials took place in the same maze, we asked participants to fill in the memorization questionnaire only once after the second, and once after the third trial. The memorization score is binary for each participant (0 = wrong choice, 1 = correct choice). Therefore, we computed Pearson's chi-squared test to detect significant associations between orientation-supporting tools and answers to the memorization question. In the first memory assessment (after the *naïve search task*) there was a significant association between the two variables ($\chi^2(3) = 27.45$, $p < .001$, $\varphi = .57$). In the second memorization assessment (after the *informed search task*) we could not find any significant relationship ($\chi^2(3) = 5.85$, $p = .054$).

Figure 4.19: Memorability assessment. Percentage of participants choosing the correct ground plan in the memorability questionnaire after completing trial 2 and trial 3.

As shown in Figure 4.19, in the first memory assessment, more participants were able to choose the right ground plan in the *None* ($n = 7$) and *Minimap* ($n = 9$) conditions than in the *Trail* ($n = 2$) and *Heatmap* ($n = 1$) conditions. In the second assessment, the number of participants who could recall the environment also differed: In the *None* condition ten participants chose the right ground plan, in the

Minimap condition seven participants succeeded ($n = 7$) and in the other two conditions considerably fewer subjects gave the correct answer (*Trail*: $n = 4$, *Heatmap*: $n = 3$).

Task Load & Difficulty After each task, participants completed the NASA Task Load Index (TLX), assessing the subjective task load. Statistical analysis showed no significant differences between orientation-supporting tools for any task. In the final questionnaire participants rated the difficulty of the given task for each of the first three trials. In all cases data were normally distributed. Significant differences only emerged in the naïve search task (*Maze Exploration*: $F(3,44) = .86$, $p = .471$; *Naïve Search*: $F(3,44) = 9.78$, $p < .001$, $\eta^2 = .40$; *Informed Search*: $F(3,44) = 2.06$, $p = .120$). In the naïve search task, participants rated the task as more difficult when no supporting tool was available (*None*), compared to the conditions *Trail* and *Minimap*. Additionally, participants in the *Heatmap* condition found the task more difficult than in the *Minimap* condition.

Subjective User Ratings In a final questionnaire, participants were asked to assess themselves. Among other things, they assessed their general ability to orient themselves and their ability to get an overview of the mazes they had explored. The questionnaire items were: (I) *How good was your orientation in the first three maze scenarios?* (1 = *very bad* to 5 = *very good*) and (II) *Was it easy/hard for you to keep the overview of the scenario?* (1 = *very hard* to 5 = *very easy*).

Significant differences were found for both ratings: (I) $F(3,44) = 6.19$, $p = .001$, $\eta^2 = .30$; (II) $F(3,44) = 4.67$, $p = .006$, $\eta^2 = .24$. Participants in the *Minimap* ($M = 3.24$) condition reported having better orientation than those in the *None* ($M = 1.76$) and *Heatmap* ($M = 1.84$) conditions. Moreover, participants in the *Minimap* ($M = 3.40$) condition rated their ability to maintain an overview on average better than those in the *None* ($M = 2.08$) condition.

Qualitative Results At the end of the study, a semi-structured interview was conducted to obtain qualitative feedback. We gathered, among others, their opinion on the helpfulness of the provided tool, strategies that were used to solve tasks, and feedback on how orientation could be improved.

Most participants stated that the orientation-supporting tool provided helped them to solve the tasks ($n = 34$; 94%). As reasons for this, three participants stated that the tool (heatmap and trail) provided outstanding characteristics that could be used as orientational anchors in an environment where everything else looked very similar and monotonous. Two participants in the minimap condition stated that the aid helped them to keep track of the maze. One participant in the *Trail* condition stated that the overlapping trails were ‘messy’ and sometimes ‘confusing’.

When asked whether they used a strategy, six participants elaborated on how they used a strategy that directly involved the orientation-supporting technique. Two used the heatmap, and one used the trail to create marks by intentionally moving in a certain pattern.

Participants were asked to estimate their performance in a real-world maze, compared to being in a virtual reality environment. Most ($n = 39$) assumed that they would have done better if they had been in a real rather than a virtual maze. The reasons they gave included the following: Teleportation hampers orientation ($n = 11$), presence of real-world artifacts ($n = 6$), movement is slower in the real world ($n = 2$), movement is easier ($n = 2$), better orientation in the real world ($n = 2$). Conversely, there were also some participants ($n = 7$) who estimated their performance to be worse in a real-world scenario due to faster physical exhaustion ($n = 4$) and the lack of teleportation for fast movement (n = 2).

As suggestions for possible improvements in orientation, participants wished for a more diverse environment that provides more distinct anchors for orientation ($n = 10$). Moreover, some participants suggested increasing the room provided for physical walking ($n = 5$). Eight participants who were not in the *Minimap* condition even asked for a minimap as additional support.

4.2.4 Discussion

For all three tasks, participants were faster when orientation-supporting tools were provided, confirming a general benefit of orientation-supporting tools, which is consistent with previous research [86, 138, 194, 318]. In exploration tasks, participants were more accurate and explored more space when provided with a *Trail* or a *Heatmap*. In the naïve search task, participants visited less space redundantly when any aid was provided. Moreover, participants with a *Minimap* were faster in tracing back to previous locations. Participants' self-assessment regarding their subjective orientation and overview capabilities in general also revealed that providing a *Minimap* was advantageous in comparison to when no orientation-supporting tool was available. Hence, we can accept hypothesis *H1* as we could show that the provision of any of the three evaluated techniques improved participants' orientation.

In the maze exploration task, participants explored more space when a *Trail* or a *Heatmap* was provided compared to when no aid was given. As a matter of fact, users in the *Trail* condition outperformed those in the *Minimap* condition in that aspect. The two techniques *Trail* and *Heatmap* performed similarly in terms of accuracy. In the maze exploration task, participants found more markers when using one of the two techniques than users without aid. If we take a closer look at the real-world inspired exploration task, the two techniques also outperformed the *Minimap*. However, between the two techniques *Trail* and *Heatmap* no differences emerged. Therefore, we can only partially accept hypothesis *H2*, since the *Heatmap* could only prevail against the *None* and *Minimap* condition in terms of exploration accuracy and efficiency.

Participants rated their subjective overview abilities as best when equipped with a *Minimap*, similar to results from a study by Kotlarek et al. [194] in which participants prefererred a 3D minimap over landmarks and waypoints. The memorability assessment revealed that participants with a *Minimap* could remember the ground plan of the maze they explored better than participants in the *Trail* condition. In the naïve search task, participants with the *Minimap* were faster compared to participants without an orientation-supporting tool or with one of the other two techniques. Therefore, we can partially accept hypothesis *H3*, stating that the provision of a *Minimap* increases the overview of users and thus promotes orientation.

In the informed search task, participants were able to navigate from the second to the third marker faster when the *Trail* was provided than if no aid was available. This could be due to the visualization of the trail that marks the path taken earlier. However, it seems that traceability skills were even further enhanced with a *Minimap*. With a *Minimap*, participants could retrace from the first to the second and from the second to the third marker faster compared to the condition in which no aid was provided. This could be due to an increased overview in the *Minimap* condition. As soon as a marker was found, participants memorized its location and could return to it more easily with the help of the *Minimap*. No significant differences emerged between the techniques. Focusing on the time participants needed to return from the last found marker to the starting position, only in the naïve search task, it took significantly less time with the *Minimap* than without any help. These findings are in line with participants' subjective assessments of their orientation and overview-keeping abilities in the final questionnaire. Therefore, we have to reject hypothesis *H4*, which is based on previous work by Ruddle et al. [318] and suggests that *Trail* is the most efficient technique for retracting tasks.

Interestingly, participants had less difficulty in remembering the maze correctly when no navigational tool was available compared to the *Trail* and *Heatmap* condition. It is possible that participants in these two conditions relied solely on the navigational aids to solve the task, which ultimately prevented participants from creating an inner image of the maze, making it more difficult for them to recall the ground plan later on. The *Minimap*, on the other hand, directly represented the ground plan of the maze, which made it relatively easy for participants to recall it afterward.

In the naïve search task, participants visited a larger area redundantly when no aid was provided. In the informed search task, only in the *Heatmap* condition, a significantly smaller area was visited redundantly by participants compared to the condition in which no orientation-supporting tool was available. Participants with a *Minimap* visited more areas redundantly compared to participants in the *Trail* or *Heatmap* condition. This could be an indication that the provision of a local overview of movement history as provided by the *Trail* and *Heatmap* techniques is more important for space-efficient, non-redundant navigation than a global overview of the environment as provided by a *Minimap*. Similarly, the task completion time was higher when participants used the *Minimap* for the naïve search task compared to participants using *Trail* or *Heatmap*. This could be due to changed user behaviour when the overview of the entire environment is given and the amount of hidden markers is known. Participants probably tried to quickly navigate through the maze to find all three markers, rather than carefully inspecting each corner as in the exploration tasks, which may have led them to frequently circle back to areas they were close by earlier.

The *Minimap* performed worse in terms of user accuracy in the real-world inspired scenario compared to the maze exploration task. In the real-world inspired task, participants in the *Trail* or *Heatmap* conditions collected significantly more markers. This could be due to the limited scalability of the minimap approach. With an increasing area of the environment, the minimap is zoomed out in order to fit the entire environment on the handheld map. As a result, everything depicted becomes smaller and harder to read.

In sum, results suggest that the deployment of orientation-supporting tools can be beneficial in exploratory search and wayfinding tasks. Depending on the task, different orientation-supporting tools are recommendable. In case retracing is a central element of the task at hand, our results suggest the use of a *Minimap*. In order to avoid redundant visits of the same location, the direct visualization of visited places by means of *Heatmap* and *Trail* visualizations is recommended. With the *Trail* visualization, participants were able to cover more of the available area on average, leaving less space unexplored. One reason for this could be that the trail enabled participants to encode the history of their movements. While the *Heatmap* representation solely highlights areas in which the user was previously located, the *Trail* visualization also provides the context of how the respective location was visited, i.e., from where and to where. Based on our results, we propose that future systems should consider the provision of visual aids in any scenarios involving search and exploration tasks, as it can reduce disorientation and task completion times while increasing accuracy. Furthermore, our results underline that less is sometimes more, as we found distracting factors of visualization elements on memorability. Hence, when memorability is an important aspect in a given task, one should consider refraining from constantly showing visual cues that are embedded in the environment (e.g., heatmaps or trails), but rather enable switching them on and off or deploying visual aids that are detached from the environment (e.g., minimaps).

4.2.5 Limitations & Future Work

Some limitations need to be taken into account. The results of the current study only allow statements about the performance of the three selected orientation-supporting tools in a planar environment where a user walks on a plane underground. Results could be different if a three-dimensional space is to be explored, for instance, if the task is to explore an abstract 3D scatterplot. In such cases, *Heatmap* and *Minimap* may not be applicable or only in a limited form. Furthermore, we used a relatively small environment and a short search task for our study. In case the environment is much larger, different levels of scalability are to be expected from the three tools evaluated in this study. Ruddle et al. [318], for example, found that overlapping trails are a major limitation of the *Trail* technique in long exploration tasks. To overcome this, the basic form could be extended by color-coding the time on the trail, reducing the size of the trail depending on the time or even blinding out old segments of the trail. Also, for the *Minimap* challenges become apparent when the environment is much larger. In this case, it may no longer be possible to display the entire ground plan on the minimap, but only a small section of it.

In the present research, we were interested in the performance of different orientation-supporting tools when the primary movement mode is teleportation. We strive to more precisely look into the impact of teleportation in future work by comparing the performance of different orientation-supporting tools when other movement options are offered. The performance of tools could change if participants could walk naturally in the 15×15 m space or use slow transitions between the starting point and the selected teleport destination.

The evaluation of different orientation-supporting tools was specifically tailored to their use in an abstract virtual reality environment. Nonetheless, all deployed techniques can also be used in augmented reality applications. However, we assume that the type of movement used - i.e., teleportation - had a substantial influence on the performance of the techniques and could, therefore, be a topic for future research. Findings, for example, about retrace and overview capabilities of the individual techniques, could be used as a starting point for subsequent research in AR. Similarly, it would be interesting to assess the actual impact of the realisticness of an environment on orientation. In the present study, we mainly used abstract mazes that could not only have led to stronger effects as desired, but also to other effects, which would make it difficult to transfer knowledge gained from the experiment with abstract mazes to everyday use cases.

Future research should also address combinations of the presented techniques, exploring if and how trade-offs could be balanced out. For example, the *Heatmap* and *Minimap* techniques could be merged into a single one combining the advantages of both techniques. This way, the user could keep an overview and, at the same time, be aware of unexplored locations.

4.2.6 Conclusions

The present study investigated three orientation-supporting tools for various exploration and search tasks in virtual reality environments. We provided the common technique of teleportation for spatial movement and used mazes as virtual environments to promote a loss of orientation. Results indicate that the tools supported users in exploration and naïve search tasks. However, we found that providing *Trails* and *Heatmaps* had a negative effect on participants' memorization ability in informed search tasks. The provision of a *Minimap* seemed to improve users' retrace ability more than the *Trail* technique, which was unexpected as the *Trail* depicts the history of a user's movement directly in the visualization. Generally, our study suggests a benefit of orientation-supporting tools and shows the strengths and weaknesses of different approaches for different tasks.

5 Assessing Applications for Immersive Analytics

Contents

Previous work revealed various opportunities and potentially advantageous properties of immersive environments for data visualization in low-level evaluations. In particular, certain properties of immersive environments, such as improved depth perception through stereoscopic vision or enhanced memorability capabilities as a result of high levels of immersion, were demonstrated. To determine the true value of potentially beneficial properties of immersive environments, applications and techniques must be developed that attempt to exploit given properties in order to improve analyses. This chapter shifts the focus to the third strategy for assessing the applicability of VR for data visualization, which is high-level, big-picture assessment. We present two examples of the implementation of such assessments of applications where VR was deployed in visual analytics procedures. In the first part, a novel approach for the comparative analysis of 3D heatmaps in VR is presented. This part is based on our CHI paper [198], which evaluates the new technique and compares it to a conventional comparative analysis approach, highlighting the task-dependent advantages and disadvantages of using VR. The second part presents an example of deploying VR for the interactive analysis of 4D scene reconstructions. This work is based on our article in the Sensors Journal [204], in which we present a prototype for an interactive and immersive exploration of time-dependent 3D scene reconstructions.

5.1 Assessing 2D and 3D Heatmaps for Comparative Analysis: An Empirical Study

Heatmaps are a popular visualization technique that encode 2D density distributions using color or brightness. Experimental studies have shown though that both of these visual variables are inaccurate when reading and comparing numeric data values. A potential remedy might be to use 3D heatmaps by introducing height as a third dimension to encode the data. Encoding abstract data in 3D, however, poses many problems, too. To better understand this tradeoff, we conducted an empirical study ($N = 48$) to evaluate the user performance of 2D and 3D heatmaps for comparative analysis tasks. We test our conditions on a conventional 2D screen, but also in a virtual reality environment to allow for real stereoscopic vision. Our main results show that 3D heatmaps are superior in terms of error rate when reading and comparing single data items. However, for overview tasks, the well-established 2D heatmap performs better.

5.1.1 Introduction

Heatmaps are omnipresent in information visualization. They are frequently used not only as a basic module for novel applications and visualization designs [316, 331, 408], but also as a tool for presenting research results [51, 274, 393]. Heatmaps allow the analyst to quickly grasp a 2D distribution because of their capability to facilitate intuitive encoding of values by color in a 2D grid. The technique is also frequently deployed for comparison tasks, for example, to convey a temporal progression of 2D distributions. When heatmaps are used in a visual analytics pipeline to display intermediate results, visualizations have to be compared with visualizations from previous analysis steps to evaluate the improvement caused by parameter changes. Side-by-side comparisons of 2D distributions are also frequent tasks. For instance, Schreck et al. [331] deployed heatmap visualizations to compare distributions of different properties in Kohonen maps. In their application, the analyst can quickly identify attributes for which their neural network is optimized for. Besides, there are many examples in literature in which pre-post comparisons are presented. For instance, when two heatmaps are compared, one of which depicts values in their original form, and a second depicts the distribution after applying a change [101, 275, 416].

However, when comparing multiple heatmaps with each other, several issues emerge. In juxtapositioned comparisons, the analyst must locate a specific location in multiple heatmaps in order to compare individual values. This leads to a high cognitive load and a high potential error rate. Supportive methods like linking and brushing can alleviate this problem but might lead to additional issues, such as a perceptual distortion of values close to the highlight. Alternatively, multiple distributions could be joined into a single aggregation visualization, such as a difference map. Aggregations like this facilitate various tasks, such as offset extraction, but make others impossible (e.g., exact value extraction).

To overcome problems of juxtapositioned comparisons, one could extend the 2D heatmap by a third dimension, double encoding value by color and height. The resulting 3D heatmaps could then be superpositioned for comparison. This might facilitate local referencing in multiple 2D distributions and thus reduce the overall cognitive load without juxtapositioning while preserving all information of each distribution.

To date, however, a vast amount of work in the InfoVis domain has pointed against the use of 3D for abstract data. 3D is accompanied by fundamental issues such as occlusion [137], perceptual distortion [286], and the absence of a common baseline [290]. After due deliberation, however, we believe that there are good reasons why the applicability and usefulness of 3D heatmaps should be further investigated. Reason one are the shortcomings of traditional 2D heatmaps for comparative analysis, as described above. Reason two are new display technologies like VR and AR that are becoming more widespread and will necessitate us to think about proper data representations within them. It is not clear yet, in how far drawbacks of 3D data representations will persist in such environments, or if they might be reduced or even balanced out in comparison to conventional screen-based visualizations. Properties unique to VR approaches such as available degrees of freedom, cognitive immersion, or interaction possibilities have been identified as beneficial in many use cases. Among others, related research revealed benefits of VR in terms of improved spatial memory [245], learning performance [300], spatial understanding [145], the understanding of geometric models [429], and collaboration aspects [96]. Advantages of using VR to observe established visualizations could be demonstrated in several cases, for example for scatterplots [12], graph visualizations [109], and flow visualizations [122].

Towards better understanding of this new design space, we present a prototype for the interactive exploration of data distributions with 3D heatmaps. In addition to standard functionality for exploratory analysis, our prototype specifically supports comparative analysis tasks on multiple heatmaps. To do so, we employ a novel interaction metaphor, where users can shift stacked heatmaps into each other for spatial and numerical comparison. In a quantitative user study we focus on the prototype's capabilities for comparative analysis and compare our visualization approach of stacked 3D heatmaps with the conventional approach of juxtapositioned 2D heatmaps. Based on a literature review, we identified the most common types of tasks in comparative heatmap analysis, including *Lookup-Tasks*, *Locate-Tasks*, and *Overview-Tasks*. In addition to comparing the two different visualizations, we also tested the type of *Medium* in our experiment. Half of the participants conducted the study on a conventional monitor screen, the other half in a virtual reality environment, additionally allowing us to assess the impact of VR on the analytic performance of users.

In summary, we make the following two main contributions: *(i)* we present a 3D heatmap prototype in VR that supports the comparative analysis of multiple distributions, and *(ii)* based on this prototype, we present the results of an empirical study comparing the performance of 2D and 3D heatmaps for comparison tasks in virtual and conventional screen environments.

5.1.2 Related Work

In the following, we will first provide a brief overview of where and how heatmaps were used in previous works. Next, we will summarize motives for deploying virtual reality as a medium for the observation of visualizations. Subsequently, we survey several existing approaches to comparative visualizations, with a focus on heatmaps.

5.1.2.1 Heatmap Visualizations

Heatmaps are a well-known technique to visualize continuous data. Their applicability and usefulness has been demonstrated in various domains, for example, in medicine for volume surface visualizations [367], in geography for temperature visualizations [89] or even for abstract trajectory analysis [331]. Often, heatmaps are used for the presentation of 2D distributions, which are the result of statistic evaluations [90, 197]. They are also frequently deployed for lining up and comparing two or more results, such as different experimental conditions or pre-post comparisons [189, 296, 331]. For the comparative analysis of 2D heatmaps, there is a large number of different techniques for merging two or more heatmaps into a blended view of them. For instance, Jo et al. [180] present various approaches to visualizing two density maps in one visualization using different blending techniques.
Three-dimensional heatmaps, also referred to as heightmaps, extend 2D heatmap visualizations by double-encoding the "heat" by a position as well (i.e., height). Of course, color can be replaced entirely or used to encode an additional attribute. Most commonly, heightmaps are associated with geographic visualizations such as OpenSpace [273] or Google Earth [135] in which landscape elevations are mapped to height. However, 3D heatmaps have also been deployed in a variety of visualizations for more abstract data, such as in sound analysis for frequency visualizations [316] or in medicine for the analysis of vascular movements [408]. Büschel et al. [46] used heightmaps to investigate spatial interaction in AR environments on 3D data visualizations. Tory et al. [369] present an empirical study for a search and value extraction task on scatterplots and 3D data landscapes. For the investigated task, the point-based spatialization was superior compared to the 3D heatmap-like representation. Our work adds to the line of work of empirically studying heatmaps, by focusing on the aspects of 2D vs. 3D heatmaps and the impact of stereoscopic perception and immersion on performance.

5.1.2.2 3D Visualizations and Virtual Reality

In general, 3D visualizations are not the preferred solution for abstract data as they are accompanied by flaws like occlusion and perceptual distortion [258]. Sedlmair et al. [335] compared the performance of 2D and 3D scatterplots for cluster verification tasks on dimensionality reduced data and, based on their results, strongly advise against using 3D visualizations for this task. However, in the recent past, virtual reality devices, such as Oculus Rift or HTC Vive, have gained attention in the field of information visualization. Dwyer et al. coined the term 'Immersive Analytics' by defining it as "the use of engaging, embodied analysis tools to support data understanding and decision making" [102]. In their book, Marriott et al. [219] provide a collection of papers that characterize this research area. Among others, they point out that immersive analytics can be an opportunity for decision making and knowledge generation even for abstract data. The deployment of virtual reality environments for 3D visualizations has proven advantageous in some cases [141, 294, 429]. According to Donalek et al. [96], improved depth perception in VR leads to a better overall perception of the datascape geometry

and a better understanding of the data in graph visualizations. Similarly, Erra et al. [109] found a beneficial effect of VR on graph exploration tasks. Etemadpour et al. [111] conducted several studies comparing stereoscopic visualizations to projection-based ones. They found that surface-based visual encoding benefits more from a VR setting than point-based renderings.
We add to this line of research by assessing design factors that have not been studied so far. For comparative tasks, superpositioned 3D heatmaps could pose a benefit compared to conventional juxtapositioned heatmaps when inspected in VR .

5.1.2.3 Comparative Visualizations

The comparison of two or more data sets is a frequent task in visual analysis. Hence, a vast amount of different visualization techniques and approaches exist for comparing several visualizations. Most commonly applied are side-by-side visualizations [170, 331]. With this approach, the observer has to find the same position in each visualization, which can be a tedious and inaccurate task. Linking and brushing can be deployed to ease this process. For time-series comparisons, Gleicher et al. [132] presented several strategies. Besides juxtaposition and superposition, the signals can be merged by calculating a difference signal and displaying it instead. However, each of the named strategies has its benefits and drawbacks. Alabi et al. [1] surveyed various techniques to compare surface visualizations. Among others, they listed the usage of transparency in combination with overlapping surfaces, the partitioning of the surfaces into slices aligned in alternating order, and the utilization of semi-opaque textures. Multiple coordinated views can be used to look at one data set from different perspectives. This design consists of different windows in which different projections or visualizations of the same data entity are displayed [37, 164, 268]. For comparative analysis, any comparison technique can be used separately in each coordinated view.

In our work, we add to this strain of research by assessing the performance of superpositioned 3D heatmaps for typical comparative tasks. Moreover, we assess the impact of immersion on user performance by deploying the compared types of visualization on a conventional screen and in a virtual reality environment.

5.1.3 Prototype Description

In this section, we will first discuss some general design considerations for visualizing 3D heatmaps in VR , and how we implemented them in our prototype. Subsequently, we will focus on comparing heatmap visualizations and explore the possible advantages of 3D and VR . Finally, we will investigate how the interactive embedding of 2D visualizations in the 3D design space can help the user to overcome the disadvantages of 3D representations. In the following section we will then focus on evaluating one aspect of the prototype in more detail: its capabilities for comparative analysis.

5.1.3.1 Design Considerations

In 3D heatmaps, the third dimension can be used to double encode the value by color and height. This strengthens the encoding since the value is additionally encoded by the more powerful visual variable "size" [22]. However, this comes at a price: drawbacks caused by the nature of 3D visualizations such as occlusion and perceptual distortion appear on the scene. Previous research has demonstrated a potential benefit of VR in various immersive analytics use cases such as improved spatial memory, more natural interaction capabilities, and better depth perception [96, 141, 429]. In order to compensate for disadvantages caused by the three-dimensionality of the visualization, we deploy the visualization in a virtual reality environment. Concerning comparative tasks, the three-dimensional visualization has the advantage that superpositioning is possible. This could ease spatial referencing and reduce the mental workload compared to juxtapositioned 2D visualizations. To follow up on this presumption, we strive to evaluate the performance of 3D heatmap visualizations for comparative analysis tasks in virtual reality environments. We assume that the three-dimensional visualization, in combination with stereoscopic vision, could be advantageous for such tasks. Therefore, we developed a prototype that provides a platform for 3D heatmap visualizations and the associated functionality for explorative analysis. In the following, we present the three most important design considerations for our prototype:

1. To facilitate an interactive visual exploration workflow, we follow Shneiderman's well-established "information seeking mantra" [339]: overview first, zoom and filter, and details on demand. A 3D heatmap visualization is used as a base visualization that provides the user with an overview of the distribution. The visualization environment must supply the functionality to adapt the visual encoding (color coding) and the representation of data (sampling rate, normalization). Furthermore, it must provide the functionality to enlarge areas of interest (zoom) and to filter data. It must also provide the functionality to extract exact value information for points of interest.

2. The prototype must provide the functionality to display multiple visualizations simultaneously, thus enabling comparative analysis tasks. We decided to align the 3D heatmaps horizontally in order to obtain a common plane of reference and, at the same time, provide the metaphor of natural landscapes. Individual layers of heatmaps should be movable only in the vertical direction, preserving spatial referencing. Shifting could help users to identify connected surfaces and partially overcome problems associated with occlusion. For instance, one heatmap can be shifted through the other until its surface pierces through the surface of the other at a specific location of interest. The extent of the required shift then indicates the offset of values at the given location.

3. Although 3D visualizations have several disadvantages, they also offer new design possibilities and metaphors. In order to take advantage of 2D visualizations, the visualization design space should not be limited to 3D. Users should be able to seamlessly transition data visualized in the 3D representation into 2D projections of the same data. The user must be able to create 2D aggregations from the 3D representation and display them in the visualization environment in order to overcome drawbacks associated with 3D visualizations.

5.1.3.2 Prototype: Base Visualization

The visualization environment was developed using the gaming engine Unity3D [371] and consists of a 5m×5m room with a table in the center, surrounded by white walls. The 3D heatmap visualization is placed on top of the table and framed by axes at each corner. The visualization can be inspected through an HTC Vive Pro [161] HMD by walking around the table.

As depicted in Figure 5.1 (top), the visual encoding can be adapted interactively. The user can change the coloring of the heatmap, emphasizing certain data ranges. Further, it is possible to select data ranges in the heatmap by hovering over an arbitrary axis and zoom into the selected data range (see Figure 5.1, bottom). Details on certain locations can be obtained by clicking on it, which opens an information pop-up.

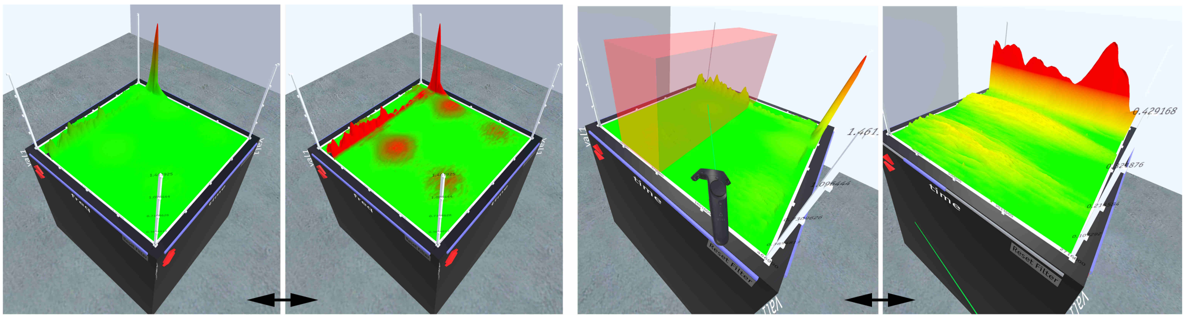

Figure 5.1: The user can interactively change the color encoding (left). Data ranges can be selected and applied as a filter to zoom the visualization to the selected data range (right).

5.1.3.3 Comparison of Multiple Heatmaps

Comparing two visualizations to find correlations or other coherences is a frequent task [132]. While conceivable, directly comparing two heatmaps with one using color and the other using opacity to encode the values is perceptually extremely challenging and not feasible in practice. Thus, in traditional visualization environments, side-by-side or selective, confined comparisons are state of the art. Another option is to merge the data sets of both visualizations, which are to be compared into a new visualization, e.g., by creating difference maps or aggregated views. Juxtapositioned small multiples increase the cognitive load in comparative tasks since the user has to coordinate his or her attention between two or more different visualizations. I.e., a particular location in one heatmap has to be located in a second heatmap in order to compare the values. Aggregated visualizations that simultaneously encode information of two or more distributions often lack crucial information. For instance, when creating a difference map of two heatmaps, only the value offset is displayed, whereas the absolute level of values disappears. Of course, these views can be displayed additionally to juxtapositioned small multiples, but this further increases the overall mental workload. Interaction methods, such as linking and brushing, can reduce the mental effort, but the fundamental challenges remain.

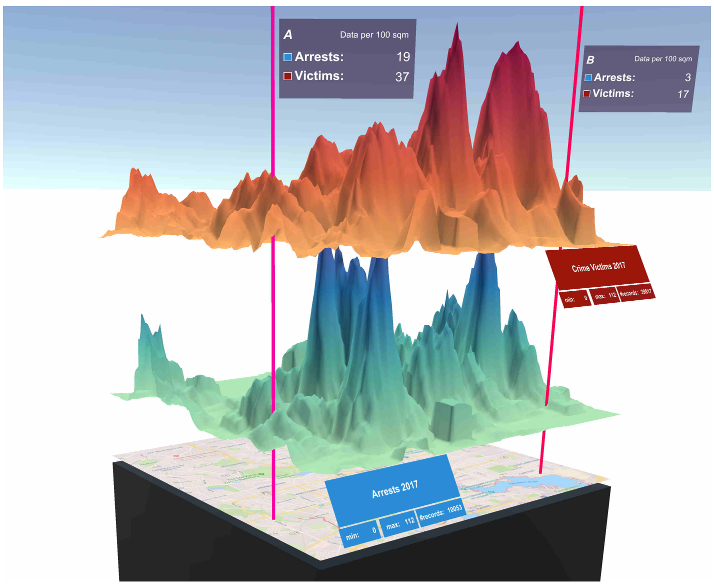

Figure 5.2: Stacked 3D heatmaps for comparative tasks. In this example, one heatmap (blue) shows the frequency of arrests at any location in a district of Baltimore. The second heatmap (red) depicts locations where crimes were reported by victims. The analyst can select any location on the map to get detailed information about that location (pillars).

Stacked 3D heatmaps in VR naturally lean themselves towards interaction techniques that allow intuitive comparisons of two or more heatmaps. Instead of coordinating attention between several visualizations, as it is the case with small multiple visualizations, we propose to vertically shift heatmaps into each other. In doing so, values can be compared along the vertical axis through a specific point of interest. Figure 5.2 depicts two horizontally aligned heatmaps stacked on top of each other. For optimal visual differentiation, each displayed 3D heatmap should have a distinct color scale. Each value in the heatmaps is double encoded by value and height, where color and height can be used to compare values within the same heatmap, and height can be used for efficiently compare values of two heatmaps. The user can shift the heatmaps interactively along the vertical axis. Shifting one heatmap relative to another eases the detection of coherence between them when peaks of one heatmap appear and rise through the other (see Figure 5.3). For exact value comparison, heatmaps can be snapped in with aligning axes to establish a common baseline for the visualizations.

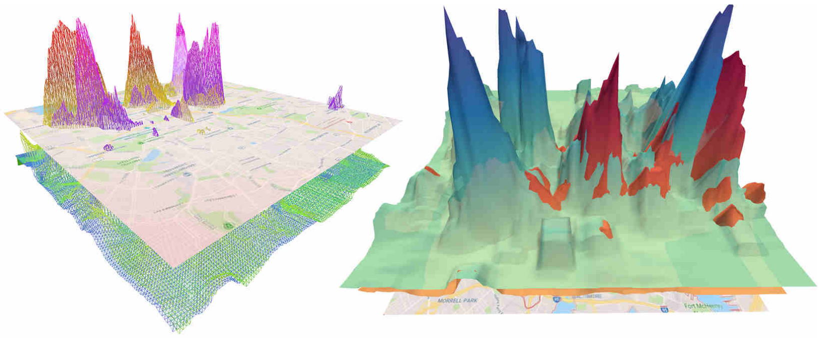

Figure 5.3: Superimposed heatmaps, one encoding arrests and the other reported crimes. Meshing (left) or transparency (right) can be used to overcome occlusion. The geo-map layer can be used as a cutting plane to serve as a common plane of reference for shifted peaks or to cover distracting parts of the visualization (left).

Besides shifting the heatmaps and the base map arbitrarily, the prototype provides further parameterization options: As shown in Figure 5.3, a base geo-map can provide further spatial reference (left), and the appearance of the heatmap itself can be customized to display a meshed surface (left) or a semi-transparent surface (right). As these representations allow users to look through surfaces, the user gets the possibility to see how strongly the values of the different heatmaps correlate in certain areas. The color maps can be adjusted interactively. For better readability and comparability, a user can interactively place labels anywhere on the map to compare values of one or more different points between all displayed heatmaps (see Figure 5.2, labels A and B). The pink lines perpendicular to the map support the comparison as visual cues, pointing out the selected position on all layers.

5.1.3.4 Transformation & Projection from 3D to 2D

Three-dimensional visualizations can have some disadvantages in visual analytics tasks. For example, occlusion and perspective distortion can occur. To overcome problems arising from 3D visualizations, we integrated the possibility to create 2D projections from selections in the 3D visualization seamlessly. The user can select the data to be projected using the selection box tool (see Figure 5.4). By pressing and holding the trigger button on one of the sides of the selection box, an aggregation is generated and attached to the controller. It can then be placed on a wall by releasing the button at the desired position. The side of the selection box that was clicked determines by which axis the data is aggregated. As if the selection box had been compressed, only the selected side remains. Projections can be arbitrarily organized, deleted, or supplemented by annotations (drawing function). For the aggregation, currently, the third dimension that needs to be reduced is averaged to demonstrate the concept. Of course, any other aggregation function can also be used.

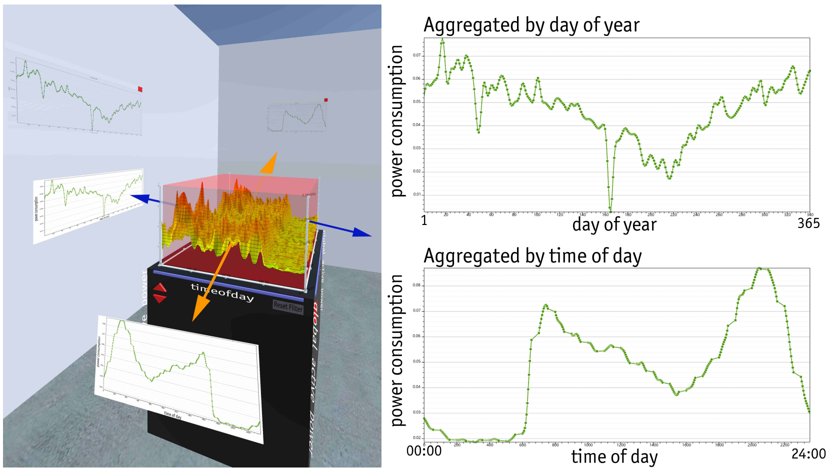

Figure 5.4: 2D projections of 3D heatmap. In this example, all data is selected (red selection box, left). Two 2D line charts (right) were created interactively by selecting two different sides of the selection box, resulting in different forms of aggregation. Example: single household power consumption over one year.

5.1.4 Study: Heatmaps for comparative analysis

In order to evaluate the previously presented approach of stacked 3D heatmaps, we conducted a user study with 48 participants. We compared the novel approach with the conventional analog approach of juxtapositioned 2D heatmaps. To assess the role of VR , we also considered the 2D version in a virtual reality environment.

5.1.4.1 Study Design

The conducted study comprises two experimental factors: *Medium* (Screen, VR) and *Dimensionality* of the visualization (2D, 3D – see Figure 5.5). We used a between-subjects design to avoid learning effects. The sample was divided into four groups, each of which is a combination of the two factors (*Screen2D*, *Screen3D*, *VR2D*, *VR3D*).

The 2D condition (Figure 5.5, left) consists of four heatmaps. The upper two heatmaps are the distributions that have to be compared. Each of the bottom two heatmaps is an aggregation of the upper two distributions. The first one (bottom left) is a difference map (values of heatmap A minus values of heatmap B), and the second one (bottom right) displays the sum of the values of both heatmaps (normalized to range between 0 and 10). The aggregation views were added to make a fair comparison to state-of-the-art methods for comparative analysis [335]. We used a color scale (blue-white-red) that is frequently deployed for comparative tasks of heatmaps with negative values [36, 51, 142, 287]. In the *VR* condition the visualization was attached to a wall standing in the virtual environment and participants were able to move in the virtual space, whereas in the *Screen* condition the 2D visualization was centered on the screen and no motion interactions were provided.

The visualization of the 3D condition is depicted in Figure 5.5 on the right. Each distribution is visualized as a meshed 3D heatmap with a uniform color for visual distinction. We abstained from using unique color gradients for each heatmap for improved visual distinctness of the two distributions.

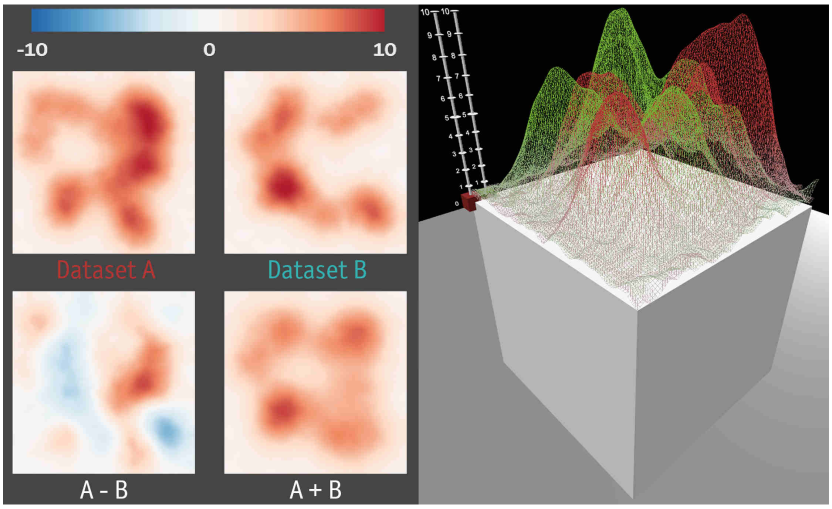

Figure 5.5: Heatmap visualizations. In the 2D condition (left), two 2D heatmaps represent the distributions to be compared (left, top). Additionally, two aggregated heatmaps show combinations of both distributions (left, bottom). In the 3D condition (right), two 3D heatmaps are superpositioned on top of a cube. Each of them has a uniform color for visual distinction.

To increase the controllability and fairness of the study, we removed advanced interaction capabilities, such as filtering, and creating 2D projections from 3D heatmaps. Participants were able to spatially navigate in the visualization environment using either keyboard and mouse (*Screen* condition) or body movements (*VR* condition). As a further interaction, it was possible to shift each heatmap up and down by dragging its red anchor cube vertically. Apart from these, there were no further interaction capabilities in the study (e.g., no zooming, no adaption of color maps).

While we sought to make the comparison between 2D and 3D conditions as fair as possible, there are some limitations stemming from the interactive nature of the design space that we intend to study. Limitations caused by these differences in interaction capabilities are discussed in the limitation section.

5.1.4.2 Comparative Tasks

For an appropriate selection of tasks, we surveyed all IEEE Vis papers of the last five years. We identified 54 papers that used heatmaps for comparative tasks and classified them into four types of tasks provided in the visualization tasks taxonomy by Brehmer and Munzner [35]: *Lookup*, *Locate*, *Browse* and *Explore*.

We did not distinguish whether the comparative task was the main focus of the paper or just implicitly mentioned in the presentation of the results. Based on the comparative tasks found, we created an abstract version of the tasks with the aim of reflecting the common purpose of each category. Due to limited controllability in the study for default *exploration* (location and target unknown) and *browsing* (location is known and target unknown) tasks, we refrained from adopting them directly. Instead, we identified an essential element from both types of tasks and merged them into a single *Overview* task. For *exploration* and *browsing* tasks, the analyst must understand the overall distribution of the heatmaps to be compared.

Lookup Target and location are known. For comparative analysis, this means that the value at a specific location has to be extracted from two different heatmaps in order to be compared with each other. For instance, Wang et al. [393] developed a visualization technique for networks in which they deployed heatmaps in the background to visualize density. They compared multiple of these heatmaps with each other by picking out a location of interest in a heatmap and reference the same location in a second heatmap. Various other examples for comparative tasks on heatmaps exist that compare values at a specific location in multiple heatmaps [51, 274, 391]. To avoid the unfair comparison between the visual variables color and height, we abstained from asking participants to extract exact values from heatmaps and created a task in which participants should estimate the distances between pairs of locations and compare the relative difference of distances. Therefore, we placed two markers in each heatmap. For each heatmap, participants should estimate the value offset and compare it relatively to the other one. Instead of asking for the total difference, participants should only indicate - with "Yes" or "No" - whether the value offset in the first heatmap is higher than the value offset in the second heatmap (see Figure 5.6). To indicate the positions to be compared, we inserted colored markers into the respective visualization. In the 2D condition, the 3D pins could be perceived as colored dots when inspected from above.

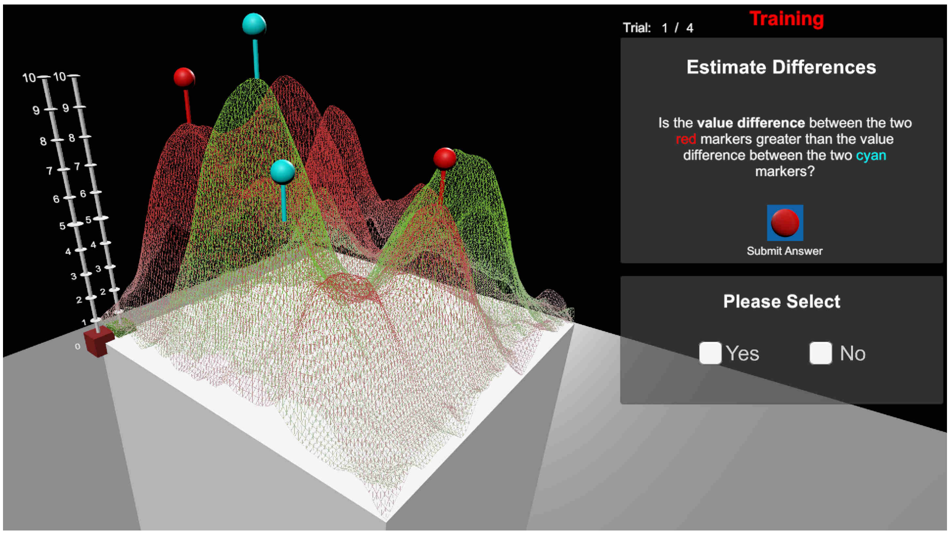

Figure 5.6: Study interface. Interaction board is attached to the right border of the screen in the *Screen* conditions and attached to the left VR controller in the *VR* conditions.

Locate Target is known, but the location is unknown. For comparative analysis, this describes a class of tasks in which the analyst visually searches for common characteristics in both heatmaps. For instance, Papadopoulos et al. [276] visualized experimental results of user movement as heatmaps. They then visually compared the heatmaps of different tasks and focused on finding common hot-spots in several heatmaps. Various analogue examples can be found in literature [2, 101, 416]. We used a task where participants had to find two locations where both heatmaps had an equally intense hot spot. Participants were asked to point out the identified shared peaks. Using the respective input device of the condition, participants could click on a heatmap to create a marker at the selected position. Markers could be re-positioned arbitrarily. There were precisely two shared peaks in each heatmap pair in all trials of this task.

Overview For many tasks of the classes *Explore* and *Browse*, it plays a vital role in keeping track of the entire distribution. Borkin et al. [28] compared pairs of eye-tracking fixation heatmaps in an exploratory manner. They did not only search for new, interesting properties in both heatmaps (*explore*), but also picked out locations of interest and investigated correlations between the two heatmaps at that position (*browse*). Many other examples of papers exist, in which such exploring and browsing tasks were applied [49, 113, 275]. In most of the tasks, an overview is a crucial factor for solving the task efficiently. Therefore, we deployed a task with which we could assess how well the overall distributions in two heatmaps can be compared. Half of the presented pairs of heatmaps were flipped and rotated versions of each other with different noise levels. So the overall structure was similar, but the overall appearance was slightly different. Participants had to judge if the second heatmap was a transformed version of the first one. Like this, we could assess if participants were able to keep track of the overall distribution in each heatmap.

5.1.4.3 Data

In order to conduct a controlled user study, we created a set of distinct distribution pairs with certain, measurable characteristics. For each task, 14 distribution pairs were generated. For each distribution, we placed 10 to 20 Gaussian kernels randomly on a 100×100 pixel grid. Each position in the grid can contain a value between 0 and 1. The kernel size (30 - 60 pixel) and the peak value at the center of the kernel (0.5 - 1.0 value points) vary randomly within the specified ranges. In a pilot study, we experimented with different parameter settings and identified the one used as the one with the best results. Participants were able to solve roughly 50% of the tasks correctly. In the end, we added random noise to each distribution (0.0 - 0.4 value points).

Lookup As this task does not need any further constraints, distribution pairs were generated as described above. This results in pairs of distinct distributions as depicted in Figure 5.7 (left).

Locate For this task, precisely two locations in both distributions of a pair need to have the same value. Therefore, we added only 8-18 random peaks in each distribution and added two more common peaks in both of them. This guarantees that for each pair, exactly two peaks exist that are of the same height and at the same location (see Figure 5.7, center).

Overview For this task, half of the distribution pairs were not altered (see Figure 5.7, left). For each pair of the other half, one distribution was randomly generated. Its counterpart was then generated by rotating it one to three times by 90° and flipping it on the horizontal axis between each rotation with a probability of 50%. The noise was applied to each distribution separately. This results in distribution pairs with one being a mirrored and rotated version of the other as depicted in Figure 5.7 (right).

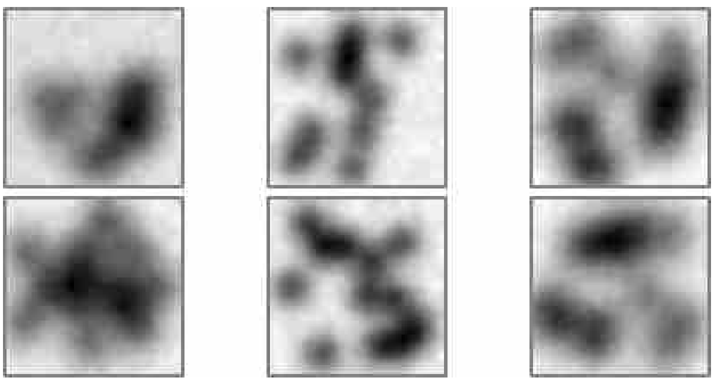

Figure 5.7: Sample distribution pairs for each task. Left: *Lookup*, center: *Locate*, right: *Overview*.

In order to be able to qualitatively assess differences in the four conditions for real-world data as well, we included an open discussion trial in which we showed real-world data to the participants and discussed it. For this, we used the Baltimore crime dataset depicted in Figure 5.2.

5.1.4.4 Procedure

Participants were welcomed and gave written informed consent. They were then introduced to the topic by reading an information sheet. The study supervisor immediately clarified questions that arose during the reading. After participants were familiar with how to interpret the base visualization (2D/3D heatmap), they were prepared to start the study trials by sitting down in front of the monitor or fitting the HMD on their heads.

The experiment was structured into three main parts. First, participants executed a block of 42 trials in total (four training trials + ten trials in each task). For each of the three tasks, they first completed four training trials in which they were introduced to the current task and confronted with the correct answer. Once they fully understood the task, participants completed ten trials without any support from the study supervisor. The order of tasks was counterbalanced (Latin Square design). The order of the deployed data was randomized (each task had a pool of datasets).

Second, participants completed ten memorization tasks. In this task, participants had to indicate for one heatmap whether the given distribution was part of the previous five trials. Half of the distributions shown were selected from the ten available candidates, and the other half were new distributions that were not shown in any of the previous trials.

Third, a real-world crime dataset was displayed and discussed with the participants. While viewing the visualization, the study supervisor explained the dataset. We showed two distributions, one depicting arrests in the city of Baltimore and the other showing reports of crimes. Subsequently, the study investigator asked several questions to determine if the visualization was well understood. Next, an open discussion on the situation in Baltimore was initiated.

After these three main parts, participants were asked to fill in three questionnaires: NASA Task Load Index (TLX), System Usability Scale (SUS), and a custom questionnaire. Finally, they were compensated for participating (10 Euros).

5.1.4.5 Apparatus

The experiment took place in laboratories at the University of Konstanz and the University of Stuttgart. In addition to the participant, a study supervisor was present in each session. Participants in the *Screen* condition sat in front of a 24" monitor with a resolution of 1920×1200 pixel. In this condition, participants interacted with the study software using a mouse and a keyboard. In the *VR* condition, participants were equipped with an HTC Vive Pro [161] and two Vive controllers.

5.1.4.6 Sample

A sample of $N = 48$ participants (28 female, 20 male) was recruited via invitations on social media channels, mailing lists, and flyers distributed around the universities. Most of the participants did not have much experience with virtual realities ($Mdn = 2$), heatmaps ($Mdn = 1.5$), and information visualization ($Mdn = 2$). Medians represent experience ratings of users on a scale from *very few* = 1 to *very much* = 5.

5.1.4.7 Dependent Variables

For each task (*Lookup*, *Locate*, *Overview*), we assessed the error rate and the task completion time. The error rate for the tasks *Lookup* and *Overview* was calculated as the percentage of incorrectly answered trials. For the *Locate* task, we calculated for each participant whether both markers were set within a small radius (10 cm) around the ground truth position of the shared peaks. The error rate was then calculated similarly to the other two tasks. The task completion time was measured as the time between two button clicks (display of visualization and login of the answer).
We assessed participants' capability to recall distributions from the last five trials. For the evaluation, we calculated a memorization rate as the percentage of correctly selected answers.
The NASA Raw Task Load Index (TLX [148]) was used to assess users' overall task load in the respective constellation of *Medium* and *Dimensionality*. Besides, participants completed the System Usability Scale (SUS [38]) to provide feedback on each condition.
Additionally, participants filled in a custom questionnaire assessing their subjective opinion. For instance, the perceived difficulty of each task, the certainty of participants' answers and the level of perceived immersion.
For the real-world data discussion, the study supervisor took notes during the conversation. Additionally, we recorded the entire conversation to encode it after the study in a video analysis procedure. We filtered out where participants had difficulties interpreting the visualizations correctly or comparing the two distributions. Moreover, we summarized which aspects participants mentioned as drawbacks and benefits of the particular condition.

5.1.4.8 Hypotheses

Based on observations from a pilot study, initial user feedback on our base visualization environment, and related literature, we derived the following hypotheses. All hypotheses were tailored to the conducted study, but could partially be extended to a broader scope.

H1 *Lookup-Task*: We expect participants to perform better with regard to reading values in the 3D condition, due to a more meaningful encoding of values by the variable height instead of color [22]. We further expect that VR poses a benefit compared to the screen representation because of stereoscopic vision due to previous research findings. Hibbard et al. [155], for instance, attributed better depth estimation and improved appreciation of 3D shapes and positions of objects to stereoscopic perception. We, therefore, expect the *VR3D* group to perform best in this task.

H2 *Locate-Task*: We expect lower error rates and task completion times in the 3D condition because of users' improved capability to detect highly granular value changes due to the deployment of the visual variable height instead of color [22]. Moreover, due to superposition, no mental mapping from one heatmap to the other is required in the 3D condition, which should also be reflected in higher performance.

H3 *Overview-Task*: We expect higher performance in the 2D condition due to juxtapositioned visualizations. Etempadpour et al. [111] identified the loss of overview in virtual environments as a critical issue. In side-by-side views, each heatmap can be observed separately while overplotting in the 3D condition hampers the perception of individual structures.

H4 *Memorability*: We expect participants to perform better in the *VR* condition due to increased spatial memory [67]. The spatial component supports the memorization of outstanding features. In VR , all heatmaps and their components are related to a physical location (immersion), whereas on the screen, no direct mapping is established.

5.1.4.9 Results

We report significant results of our quantitative analysis as well as qualitative feedback from the real-world data discussion. All statistical tests were performed using IBM SPSS Statistics (version 25) and are based on a significance level of $\alpha = .05$. For each dependent variable, we first tested whether the data was normally distributed (Kolmogorov-Smirnov). Depending on the outcome, we used either a one-way independent ANOVA for normally distributed data or its non-parametric counterpart, the Kruskal-Wallis test. Task load (TLX) and usability (SUS) were the only two dependent variables with normally distributed data. As post hoc tests, we deployed the Tukey-HSD test or the Mann-Whitney test (non-parametric). All info graphics depict mean values with error bars indicating the standard error of the mean. Asterisks indicate significant differences (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

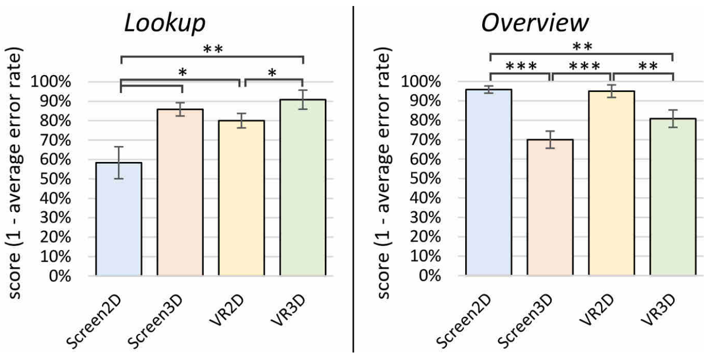

Figure 5.8: Average user performance scores in the four conditions. The score is calculated as 1 - error rate, where the error rate is the percentage of incorrect answers. The score reflects user performance. Only significant results are reported.

Error Rate Error rates differed significantly between the four investigated conditions for the tasks *Lookup* ($H(3) = 13.62, p = .003$) and *Overview* ($H(3) = 23.88, p < .001$). Figure 5.8 depicts pairwise comparisons between the four conditions. In the *Lookup* task, users scored significantly lower in the *Screen2D* condition ($Mdn = 0.65$) compared to all other conditions: *Screen3D* ($Mdn = 0.90, U = 29.50, z = -2.48, p = .013, r = -.36$), *VR2D* ($Mdn = 0.85, U = 38.00, z = -1.99, p = .047, r = -.29$), and *VR3D* ($Mdn = 1.00, U = 102, z = -2.89, p = .004, r = -.42$). Additionally, for the VR conditions, participants in the *VR2D* condition performed worse than in the *VR3D* condition ($U = 108, z = -2.51, p = .012, r = -.36$). Thus, except for the pairwise comparison between *Screen3D* and *VR2D*, participants performed worse in 2D conditions compared to 3D conditions.

For the task *Overview*, participants performed significantly better in the 2D conditions compared to both 3D conditions: *Screen2D* ($Mdn = 1.00$) lead to better results than *Screen3D* ($Mdn = 0.75, U = 7.00, z = -3.86, p < .001, r = -.56$) and *VR3D* ($Mdn = 0.80, U = 103.50, z = -2.82, p = .005, r = -.41$). Similarly, the performance was better in the *VR2D* ($Mdn = 1.00$) condition than in the *Screen3D*

($U = 88.5, z = -3.67, p < .001, r = -.53$) and *VR3D* conditions ($U = 105, z = -2.75, p = .006,$ $r = -.40$).

When comparing conditions solely based on the independent variable *Dimensionality*, for all three tasks differences emerge (see Figure 5.9): *Lookup-Task* ($H(1) = 9.96, p = .002$), *Locate-Task* ($H(1) = 7.12, p = .008$), and *Overview-Task* ($H(1) = 21.92, p < .001$).

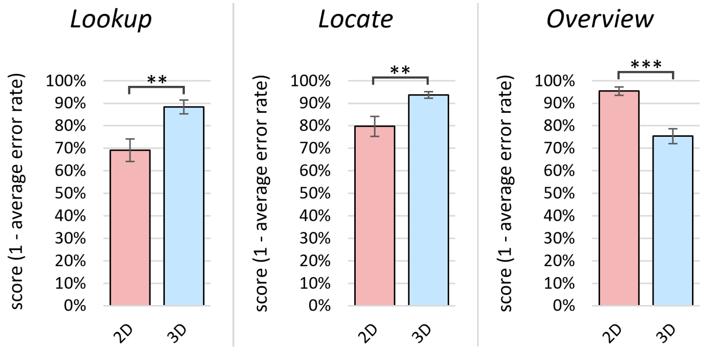

Figure 5.9: Average user performance scores by *Dimensionality*. The score is calculated as 1 - error rate, where the error rate is the percentage of incorrect answers. The score reflects user performance.

Task Completion Time As depicted in Figure 5.10, the tasks *Lookup* ($H(3) = 8.52, p = .036$) and *Overview* ($H(3) = 24.90, p < .001$) revealed significant differences between groups. In the *Lookup* task, participants required more time in the *Screen3D* condition ($Mdn = 21.51$s) compared to both VR conditions: *VR2D* ($Mdn = 9.56$s, $U = 23.00, z = -2.83, p = .005, r = -.41$), and *VR3D* ($Mdn = 11.69$s, $U = 38.00, z = -1.96, p = .050, r = -.28$).

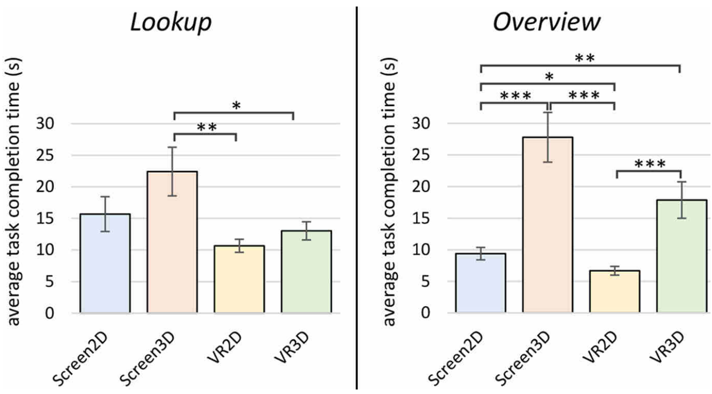

Figure 5.10: Average task completion times in seconds per task and condition. Only significant results are reported.

In the *Overview* task, participants required significantly more time in both 3D conditions compared to 2D conditions. They were faster in the *Screen2D* condition ($Mdn = 9.04$s) compared to the *Screen3D* condition ($Mdn = 30.42$s, $U = 17.00, z = -3.18, p = .001, r = -.46$) and in the *VR2D* condition ($Mdn = 6.24$s) compared to the *VR3D* condition ($Mdn = 15.79$s, $U = 8.00, z = -3.70, p < .001, r = -.53$). When comparing the two 2D conditions, participants performed faster in VR ($U = 38.00, z = -1.96, p = .050, r = -.28$).

Memorization With regard to the calculated memorization score, no significant differences emerged between groups or single variables (*Medium*, *Dimensionality*).

Task Load & Usability Figure 5.11 depicts the results of the NASA TLX questionnaire and the SUS questionnaire. For the task load an overall difference between groups could be detected ($F(3,44) = 3.67, p = .019, \omega = 0.38$). Post hoc tests revealed only one significant difference between single groups: the condition *Screen2D* ($M = 29.51, SD = 16.35$) was perceived as less demanding than the condition *VR3D* ($M = 49.31, SD = 13.25$).

In addition, the groups differed in terms of assigned usability scores ($F(3,44) = 3.99, p = .013, \omega = 2.91$). Post hoc tests revealed a significant difference between the two VR conditions. Participants evaluated the *VR2D* condition ($M = 32.33, SD = 5.71$) with higher usability scores than the *VR3D* condition ($M = 24.17, SD = 7.59$).

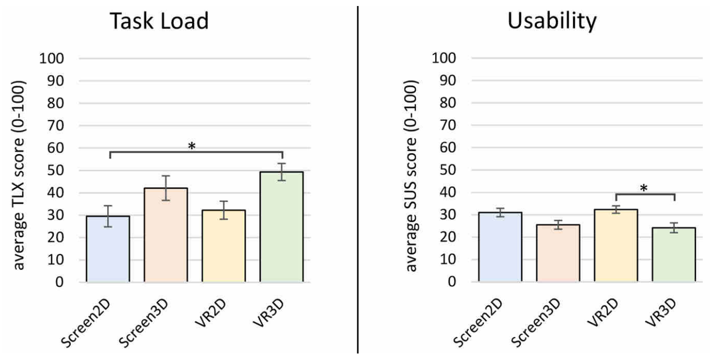

Figure 5.11: Left: average task load scores of the TLX Questionnaire. High values indicate a high subjectively perceived task load. Right: average usability scores of the SUS Questionnaire. High values indicate high perceived system usability.

Qualitative Feedback Throughout the real-world data discussion, various statements were frequently made by participants. Most of them referred to the *Dimensionality* of the visualization. In the *2D* condition, several participants mentioned that the comparison was hampered due to the requirement of switching between two different visualizations and finding one position in the two heatmaps ($n = 5$). Moreover, it was difficult for them to make out small value changes from the linear color gradient ($n = 4$). Opinions differed regarding the usefulness of aggregation views: Two participants said that they were particularly helpful and two said they saw no benefit in them.
For the *3D* condition, participants indicated the low resolution of the surface grid ($n = 4$) as a hindering factor. Two mentioned that the capability to shift one layer into the other eases the distinction between the two layers and the detection of commonalities and differences. Participants also mentioned occlusion and overlap as factors that limited their overall performance. In the *Screen3D* condition, participants found the interaction with keyboard and mouse unfamiliar ($n = 4$). In the case of the *VR3D* condition, two participants emphasized that stereoscopic vision was advantageous to observe the 3D structure of the heatmaps, to distinguish them, and find differences such as common peaks.

5.1.5 Discussion

The *Lookup* task required the user to compare pairs of values extracted from heatmaps. Results show that for each medium (*Screen* and *VR*) participants had lower error rates in the *3D* condition. With regard to the task completion time, no significant differences emerged. However, a reversed trend appeared, which is reflected in higher median task completion times in the *3D* condition. This could be due to increased interaction effort with more degrees of freedom to find the optimal perspective on the visualization. Participants' statements also underline their difficulty in perceiving the exact value from colors in 2D heatmaps. Based on these results, we can, therefore, accept hypothesis $H1$.

In the *Locate* task, participants should find commonalities in the comparative analysis. They were asked to scan two heatmaps for positions where both have equally high hot spots. The statistical analysis showed no significant differences between the four conditions in terms of error rate and task completion time. Thus, $H2$ has to be rejected. However, when comparing only *2D* and *3D* conditions, participants performed better with regard to the error rate when using 3D heatmaps. Statements of participants also reflect an advantage of the 3D representation of the heatmaps. The ability to shift heatmaps into each other allows the user to identify the offset between two heatmaps at any position quickly. By shifting one heatmap up and through the other, its peaks rise through the surface of the first one. This could also be helpful for similar tasks where smaller correlations between two heatmaps are of interest.

In the *Overview* task, participants required significantly more time in the *3D* conditions. Moreover, they performed worse in the *3D* conditions in terms of the error rate. Therefore, $H3$ can be accepted. This could be due to an improved overview caused by juxtapositioned heatmaps. In the side-by-side visualizations, the overall distribution can be observed more quickly, whereas perspective distortion, overlapping, and occlusion makes it hard to observe the entire shape of a distribution.

Concerning memorization scores, no significant differences emerged between conditions. $H4$ can, therefore, not be accepted. This could possibly be due to the fact that the visualization was limited to a small space in the virtual environment. Participants were not required to move around a lot. If participants were standing on the 3D heatmaps and surrounded by the visualization, they possibly could have made more use of their spatial memory for the memorization task.

Differences between the two 2D conditions (*Screen* vs. *VR*) were partially significant (e.g., task performance in the *Lookup* task). This unexpected result could be due to different actual sizes in which visualizations were perceived with the respective media. While the visualization in the *Screen* condition was limited to a 24" screen, participants could approach the virtual wall in the *VR* condition and thus perceptually enlarge the 2D visualization. Future work could pursue this finding by controlling the size of the visualization in VR .

Moreover, future work could follow up on our findings and extend the study design by an independent variable for the appearance of 3D heatmap surfaces. As Tory et al. [369] discovered for non-immersive 3D landscapes, the effectiveness of the visualization increased when double encoding values with height and color. Additionally, it would be interesting to qualitatively assess a potential benefit introduced by a hybrid visualization design space that facilitates the seamless transition between the 3D visualization and 2D projections.

5.1.6 Limitations

As with all empirical work, our study comes with limitations. Most importantly we note that the compared dimensionality conditions (2D and 3D) differ substantially in their visual representation and the interaction capabilities offered. We strove to minimize the degrees of freedom for both representations by providing only the tools necessary to complete the given tasks. In the 3D version, the possibility to shift heatmaps up and down is a crucial component of the technique itself, while, in previous research, the 2D version is commonly used without any interaction. This choice might have influenced our results. There is a variety of possible interaction techniques for 2D heatmaps that could potentially increase its performance. Future work might, for example, introduce a tool to interactively filter or select value ranges in the 2D condition. Similarly, we decided not to use double encoded 3D heatmaps. Instead, we tried to evaluate the impact of using 3D by comparing uniformly colored 3D meshes with flat 2D heatmaps. Again, a large amount of different visualization design options exists that could have had an impact on our results, such as colored and semi-transparent surfaces.

With regard to the lookup task, the way the task was set might have affected user performance. Colored markers were placed on the surface in the 3D condition and participants were asked to compare relative value offsets by solely considering the provided annotations (markers). Since we displayed the 3D heatmaps as meshes, participants were able to look at the visualization from the side, reducing the task to a vertical offset comparison task. This favors the 3D condition for this task. Therefore, choosing a different visualization design, such as double encoded surfaces, might lead to different results.

In the 2D condition, we used a blue-white-red color map that is often used for comparative tasks in heatmaps. The choice of the color map can substantially impact user performance though. Therefore, the use of other color maps, which, for example, highlight zero values more clearly, might increase the performance of users in the 2D condition.

We only assessed the comparative analysis of two heatmaps at a time. If more than two heatmaps are compared simultaneously, a matrix of heatmaps might be more scaleable than superpositioned heatmaps. Also, most of the participants did not have much experience with heatmaps. In particular, for aggregated 2D heatmaps, we expect a steep learning curve, which could increase the performance of expert users. Hardware constraints caused by the current state of technology for HMDs may also have affected the overall performance of participants in the VR setting.

5.1.7 Conclusion

We presented an approach for the comparative analysis of heatmaps. In a quantitative user study with 48 participants, we compared our approach to a common alternative of juxtapositioned 2D heatmaps. In addition to comparing the two different types of visualizations, we assessed the impact of immersion on the overall performance of users. Results of the user study indicate that for value extraction tasks and property detection tasks, the 3D approach outperforms the conventional visualization in terms of lower error rates, but requires more time. Juxtapositioned 2D heatmap visualizations, on the other hand, were providing a better overview of both distributions, allowing a better comparison on higher levels. We can conclude that 3D heatmap visualizations can indeed be a suitable representation in specific comparative analysis tasks. However, analysts should always consider the cost-benefit ratio when introducing 3D visualizations for abstract data. A possible solution, yet to investigate, is to make use of hybrid design spaces, cherry-picking benefits of both (2D and 3D) visualization design spaces.

5.2 Toward Mass Video Data Analysis: Interactive and Immersive 4D Scene Reconstruction

The technical progress in the last decades makes photo and video recording devices omnipresent. This change has a significant impact, among others, on police work. It is no longer unusual that a myriad of digital data accumulates after a criminal act, which must be reviewed by criminal investigators to collect evidence or solve the crime. This paper presents the VICTORIA Interactive 4D Scene Reconstruction and Analysis Framework ("ISRA-4D" 1.0), an approach for the visual consolidation of heterogeneous video and image data in a 3D reconstruction of the corresponding environment. First, by reconstructing the environment in which the materials were created, a shared spatial context of all available materials is established. Second, all footage is spatially and temporally registered within this 3D reconstruction. Third, a visualization of the hereby created 4D reconstruction (3D scene + time) is provided, which can be analyzed interactively. Additional information on video and image content is also extracted and displayed and can be analyzed with supporting visualizations. The presented approach facilitates the process of filtering, annotating, analyzing, and getting an overview of large amounts of multimedia material. The framework is evaluated using four case studies which demonstrate its broad applicability. Furthermore, the framework allows the user to immerse themselves in the analysis by entering the scenario in virtual reality. This feature is qualitatively evaluated by means of interviews of criminal investigators and outlines potential benefits such as improved spatial understanding and the initiation of new fields of application.

5.2.1 Introduction

Image and video footage is becoming increasingly important for criminal investigation, as more and more sensors, from security cameras to mobile phones, are easily available and in use. This has an impact on the accumulation of data that needs to be thoroughly investigated, which is often done manually and thus time-consuming and cost-intensive. The German police expects approximately 8 h of investigation time for one hour of video material [298]. In cases where the police ask citizens to upload video or image data for an incident, it is expected that several images will be uploaded, capturing the same content from different perspectives. Famous examples are the Boston Marathon Bombing and New Year's Eve at Cologne Cathedral [267]. Investigators determine the relevance for the data provided and note whether the supplied video records the scene of interest at the time of interest. A second step is spatial localization, which is used to determine the location of the camera sensor and its field of view. The third step is temporal localization, which establishes a temporal relationship between the other available video data. Finally, a detailed analysis of the video and image content is performed to identify objects, persons, and scenes of interest necessary for the particular case.

Our proposed visual analytics approach (ISRA-4D 1.0) supports the user in all these tasks. It automates processes to a great extent, while still allowing the user to intervene and optimize during all steps. The processed scene is composed of a 4D scene that combines multiple video sources synchronized on a single timeline. Users can explore this 4D scene in our interactive 4D scene investigator, tracking objects across various video feeds, annotating scenes, and exploring the scene in virtual reality, which significantly improves the perception of distances, angles, and details of the scene. In the long-running VICTORIA project (`https://www.victoria-project.eu/`), numerous internal and external stakeholders underlined the importance and necessity of such an approach for their daily work. Furthermore, additional use cases could be identified.

When massive amounts of data are available, for instance, through upload platforms asking the public to upload videos of an incident, the police is often confronted with a lot of irrelevant material. Our scene reconstruction approach can automatically determine whether specific images and videos where taken of a specific scene. Therefore, the police needs to reconstruct a static scene where the incident took place. Afterward, the reconstruction algorithms can determine whether the additionally uploaded video material fits into the scene or not. This approach is robust, as for videos multiple frames are available and thus more evidence can be gathered.

The primary use case is, however, crime scene reconstruction, where image material can be collected from witnesses in combination with image material recorded by the police after the incident. The constructed 4D scene can then be further annotated and explored using ordinary desktop computers and virtual reality with available consumer hardware. This allows persons involved in the case to better understand and orient themselves at the crime scene, even if they may not have seen it in reality. Additionally, such scenes can be digitally archived and also used in court.

Besides, the framework can be deployed for efficient monitoring of critical infrastructures and public places, such as airports, train stations, or industrial areas. The state-of-the-art uses arrays of monitors showing the live streams of cameras. Such an array of monitors is difficult to oversee, and important events can easily be overlooked. In addition, it requires a constant cognitive workload to recognize and remember position and orientation of each camera, which makes it increasingly difficult to trace moving objects. Our approach allows the embedding of cameras into the 4D scene, whereby the images can be projected into the scene in real-time. Additionally, the proposed concept can be used for mission planning and training for special forces in which virtual reality is an essential component. It allows users to spot a scene using drones, video glasses, or other imaging sources and receive a 3D scene that can be virtually inspected to plan the mission. Especially, the collaborative virtual reality and mixed interactions with desktop access that provide an overview are considered useful.

This work is a direct extension of an earlier publication [289] in which the predecessor framework is presented in less detail. In the line with this work, this publication contributes (1) a modular pipeline approach for the reconstruction of static 3D and dynamic 4D scenes, (2) a visual interface concept for the interactive and immersive exploration of such scenes, and (3) four use cases demonstrating the manifold applicability of our approach. The 4D scene reconstruction pipeline is carefully constructed to increase its robustness, extensibility, and user handling. The final reconstructed scene can be investigated on desktop computers, providing a good overview of the progression of events. Additionally, virtual reality allows the operator to immerse into the scene where distances, angles and orientation are perceived as in reality. The scene can be further annotated and investigated using various tools for spatial and temporal analysis to find interesting locations and times within the scene and timeline. Furthermore, our approach allows the operator to always intuitively access the original material that has been used to reconstruct the scene as well as the dynamic material that is blended into the scene.

5.2.2 Related Work

In this section, first, an overview of existing approaches and techniques is provided which can be used to analyze the content of videos. Subsequently, current state-of-the-art multi-video surveillance systems are presented. Finally, the use of visual and immersive analytics approaches in different domains is outlined.

5.2.2.1 Video Content Synthesis: Object Detection and Re-Identification

The generation of data about image content, such as the detection of objects in images and videos, is a common task used in numerous domains. One area that receives a lot of attention is research on real-time object detection. For example, YOLO, introduced by Redmon et al., is a framework based on neural networks enabling the detection of objects within images with little computational effort [305]. In later years, gradual improvements of the YOLO framework were presented: YOLOv2 [306] and YOLOv3 [307]. Besides, many alternative approaches for real-time object detection in videos were established, such as SSD [150] and R-FCN [81].

In addition to the mere object detection within an image, it is also essential to identify the same object during a video or in different footage. The so-called object re-identification task is a very challenging and error-prone task, e.g., due to context-related problems like occlusion, noise, varying illumination, moving background objects, and ambiguity [163, 431]. Li and Loy presented an approach that allows the re-identification of objects in successive frames, even if an object could not be identified in the frames in between [224]. Their segmentation-based approach allows visual tracking of objects that even change in scale and rotation. While this approach focuses on object re-identification in a single camera, others specialize in object tracking through multi-camera systems. For example, Bialkowski et al. [24] presented a database for the re-identification of persons with videos that record the same environment from different angles and under different lighting conditions. They demonstrated the dataset using a simple re-identification system that compares detected objects between different cameras. The presented approach requires overlapping viewports of the cameras. Other approaches are even more sophisticated and support the tracking of objects through non-overlapping camera networks [284, 420]. Beyond the extraction of movement trajectories, several approaches aim to analyze the movement of detected persons further. For example, Devanne et al. analyzed the trajectories of skeletons and focused on the recognition and classification of actions within the movement of a person [91]. Goffredo et al. dealt with gait analysis in surveillance videos [133]. The way a person walks is very individual, making it possible to use gait characteristics for person re-identification.

Depending on the choice of the object detection and re-identification approach, the run times and results vary. Thanks to the modular design of the current approach, new improvements of such models can easily be implemented in the pipeline. In the current version of the presented framework, a pretrained YOLO v3 module [307] was used in combination with a state-of-the-art re-identification approach.

5.2.2.2 Multi-Video Surveillance Systems

Another research focus is on the optimization of multi-camera surveillance systems. Here, the dominant goal is to contextualize heterogeneous video sources with different viewports, light and color differences, and structurally different parameters (e.g., camera intrinsics). For example, Collins et al. presented a framework for the seamless tracking of moving objects through a network of surveillance cameras [69]. A site model of the monitored environment and calibrated cameras are required to calculate the trajectories of objects. There are alternative approaches that do not require a spatial model of the environment with calibrated cameras, but rather estimate relative camera locations and their intrinsic parameters on the fly. For example, Javed et al. presented a large-scale surveillance system that automatically calculates the spatial relation between the cameras [176]. The system detects and tracks objects and persons across multiple cameras. First, the tracks of objects are computed for each camera. Then, a match between the views of the same object by multiple cameras is calculated. This makes it possible to find relationships between the field of view lines of different cameras without explicit camera calibration. Several approaches in literature (see, e.g., in [186, 222]) follow a similar principle for scenarios where it cannot be assumed that there is sufficient visual overlap occurs which would allow a purely visual camera correspondence estimation.

Other work deals with the quantification of camera constellations, calibrations, and image content. For instance, Zaho and Cheung presented a technique for optimizing the camera placement in a multi-camera system by measuring and comparing the performance of different camera constellations for object and face detection tasks [430]. Lim et al. suggested an approach for automatic, image-based calibration of stationary cameras [225], i.e., the automatic configuration of pan, zoom, and tilt parameters of cameras in multi-camera systems to optimize the system's overall performance. Beyond that, Shen et al. proposed an approach to quantify the content of surveillance cameras to prioritize the views of specific cameras in multi-camera surveillance systems [58].

The current framework comprises publicly available, state-of-the-art approaches for object detection and re-identification. The output of these models is used to improve dynamic point cloud generation processes and to simultaneously display high-level information from multiple videos in a shared 3D environment. With the modular design of the introduced framework, it is possible to adapt to further advances in this area by exchanging individual modules in the preprocessing pipeline and adapting their output to the required format.

5.2.2.3 Visual and Immersive Analytics

Visual analytics has proven to be a valuable tool for explorative and confirmatory analysis tasks [182, 183, 322, 415]. With the help of visualizations, hidden information in the data can be spotted without a concrete definition of a hypothesis. In contrast to merely statistical evaluations, this makes it possible to keep users up to date during interactive data analysis procedures. Various visual analytics solutions have also been developed in the field of police and law enforcement. For example, Malik et al. presented an instrument for police resource allocation and predictive analytics [238]. Various other works also deal with the identification of criminal hotspots and use visual analytics procedures to facilitate the process [32, 327]. Sacha et al. introduced a tool for the interactive analysis of spatio-temporal metadata of crime reports using abstract data visualizations such as correlation matrices and scatterplots [320]. Similarly, Jentner et al. analyzed crime reports, but focused on the analysis of patterns to provide insights on a large bulk of data and to find clusters of similar crimes [178].

Virtual reality has been frequently used for simulation [280], training [244], and educational [145] purposes due to its ability to immerse users in virtual environments. Virtual content can be observed more naturally, conveying the impression to experience a real situation. Immersive analytics [219] is a relatively new field in which visual analysis procedures are performed in immersive environments such as augmented or virtual reality environments. Previous research has identified several benefits associated with induced immersion. For example, Probst et al. used VR to explore large chemical spaces in which molecules are depicted as volume visualizations [294]. They concluded that VR provides a more intuitive exploration process, which is particularly useful for educational and training purposes. Zhang et al. found a benefit of VR in terms of understanding geometric structures in VR and attribute this effect to the natural inspection of 3D objects, which is similar to the inspection of physical objects in the real world [429]. Similar effects were also reported regarding more abstract data visualizations. For instance, Donalek et al. reported a better perception of the datascape geometry in graph visualizations when participants were immersed in VR [96]. Further benefits have been identified in terms of data validation [206], collaboration [96], increased task performance on specific data exploration tasks [198, 208], and memorability [245]. Etemadpour et al. found that especially surface-based visual encodings profited from a stereoscopic perception in VR [111].

The use of visualizations for the analysis and extraction of knowledge from data has proven itself in the past. Therefore, visualizations are used in the current framework to facilitate the analysis process of mass video data. Recent developments in immersive analytics research could demonstrate various advantages of using virtual reality in the visualization context. The current framework allows users to observe 4D scene reconstructions in VR in order to exploit these benefits.

5.2.3 Crime Scene Analysis Framework: Processing Pipeline

In order to explore heterogeneous data sources in a shared 3D reconstruction, the underlying data needs to be preprocessed. In this section, a detailed overview of the used preprocessing pipeline is provided (see Figure 5.12). First, all supported input data types and data-specific terms are introduced. Subsequently, the approaches used for static and dynamic scene reconstructions as well as metadata extraction (high-level scene analysis) are explained. The section concludes with the description of the module for temporal synchronization.

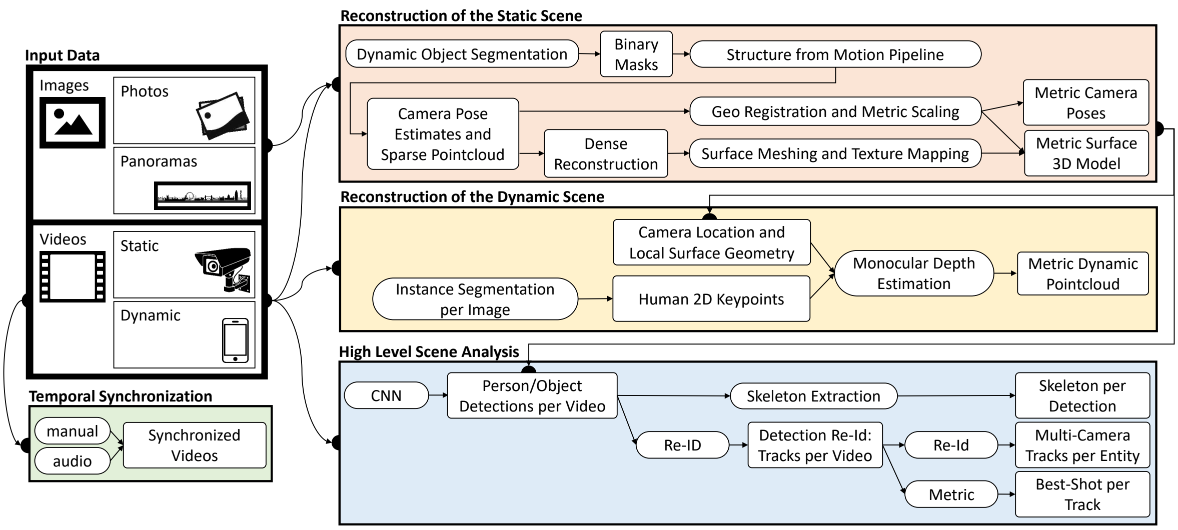

Figure 5.12: Processing pipeline of the crime scene analysis framework. Multimedia input data are processed in three main steps: First, a static reconstruction of the crime scene is created using a structure-from-motion approach. Second, dynamic elements are extracted as dynamic point clouds. Third, tracks of persons and objects are extracted using machine learning models. The 4D reconstruction pipeline was conceptualized and implemented by T. Pollok.

5.2.3.1 Input Data

The crime scene analysis framework presented is optimized for the rapid analysis of large amounts of image and video data from a certain incident. For example, after a shooting in a city center, sources could be recordings from surveillance cameras as well as photos and videos taken by eyewitnesses with their mobile phones. Therefore, the resulting set of data sources can be very unstructured and difficult to analyze. The two main sources are static cameras that do not move and maintain their perspective (static camera), and moving cameras that record different locations throughout the incident (dynamic camera). To further enhance the context, image and video material from the time before or after an incident can also be integrated into the framework by registering it solely in space, without considering time. These sources are *time-independent*, as they are not registered to a certain point in time of the progression of events to be analyzed.

Besides video data, image data such as individual photos, panoramas, and photo spheres can also included in the analysis. By default, such footage is currently treated as time-independent and is only registered in space. This could, for example, comprise images and photo spheres taken from the place of interest after the incident in the forensic analysis and help investigators to compare the environment at the time of the incident with the environment shortly after the incident.

5.2.3.2 Reconstruction of the Static Scene

The preprocessing pipeline distinguishes between the reconstruction of the static environment (static reconstruction) and dynamic contents, which are time-dependent (dynamic reconstruction). The goal of the static reconstruction is to create a digital clone of the environment of interest - for instance, of an environment in which a crime happened. The therewith created static 3D model can be used as a framework to embed additional, time-dependent elements, such as detections of persons, annotations, or dynamic reconstructions.

Figure 5.13: Sparse reconstruction (top left), created through Structure-from-Motion approach by matching key-characteristics in different source images - e.g., feature correspondence of pixels in red circle.

Figure 5.12 depicts all steps leading to a static reconstruction. First, any dynamic contents are detected and removed from the video frames (object detection and masking). Subsequently, a Structure-from-Motion approach [330] is deployed to match keypoints in different images, create a sparse pointcloud, and estimate the camera location for each frame. Figure 5.13 depicts three frames from handheld video cameras and drone recordings in which keypoint correspondences were found (red circle). The respective point is then registered to a certain location in the sparse 3D pointcloud and allows, together with many other matched keypoints, the estimation of camera intrinsics (camera parameters) and extrinsics (camera location) for each frame. As a next step, the sparse reconstruction is enhanced and transformed into a dense reconstruction using a multi-view stereo reconstruction with OpenMVS [55]. The resulting reconstruction now consists of meshed surfaces, i.e., textured triangles instead of colored points. As the last step, the static reconstruction needs to be scaled to the real-world metric system. As the automated approach does only calculate relative offsets between corresponding points, the original reconstruction is not metrically scaled. Therefore, the user needs to geo-register the reconstruction by assigning reference points in the reconstruction to locations on a geo-map or by manually declaring several sample distances in the reconstruction. After the reconstruction is geo-registered or manually scaled, distances in the reconstruction resemble real-world distances and can be easily calculated.

5.2.3.3 Reconstruction of the Dynamic Scene

The reconstruction of dynamic contents considers dynamic, frame-dependent elements. The currently deployed approach is as follows. Each time-dependent camera is first segmented into dynamic and static components for each frame. Each dynamic component classified as a human is then processed in a skeleton extraction procedure. Extracted skeletons, in combination with the previously estimated camera intrinsics and extrinsics (static reconstruction), are then used in a monocular depth estimation approach to calculate the depth of each pixel in each frame. As a result, a depth-map is created for each frame containing information on the depth of each pixel, starting from the camera's origin. Figure 5.14 depicts exemplary results created with this approach. On the left, two frames, each of which being displayed in the visual exploration tool as dynamic 3D reconstructions. For a more detailed description of the deployed reconstruction approaches (static and dynamic) - please have a look at the original paper [204].

5.2.3.4 High-Level Scene Analysis

After starting the main application, the system checks whether feature preprocessing has been performed beforehand. If not, the preprocessing sequence will be started and each *time-dependent* video is processed in an object detection pipeline. The pipeline for the high level scene analysis (see Figure 5.12) has a modular structure. This way, the entire pipeline or parts of it can be replaced by other modules that deliver an output using the same format. Figure 5.15 shows a frame from the feature extraction preprocessing step. During preprocessing, the video is played back and all recognized objects and persons are highlighted by colored rectangles, including their respective path of movement.

Object Processing in Camera Space Most of the feature extraction pipeline takes place in camera space. The position of detected objects is described in pixel coordinates, based on the frame in which they were detected. All further steps of feature processing (e.g., skeleton extraction and re-identification) are image feature-based and therefore do not make use of the position itself. In a subsequent step, which is described in the next paragraph, these pixel coordinates are mapped to world coordinates using the estimated camera pose from the reconstruction.

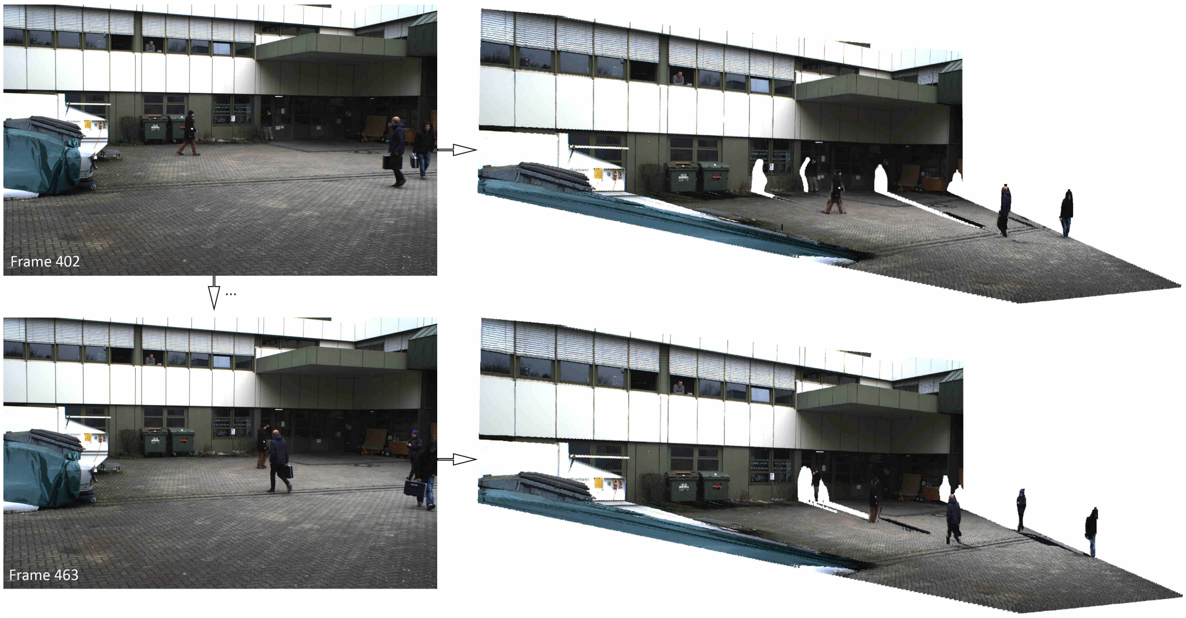

Figure 5.14: Two frames from a video (left) are depicted as 3D reconstructions with estimated depth information in the 4D scene exploration tool.

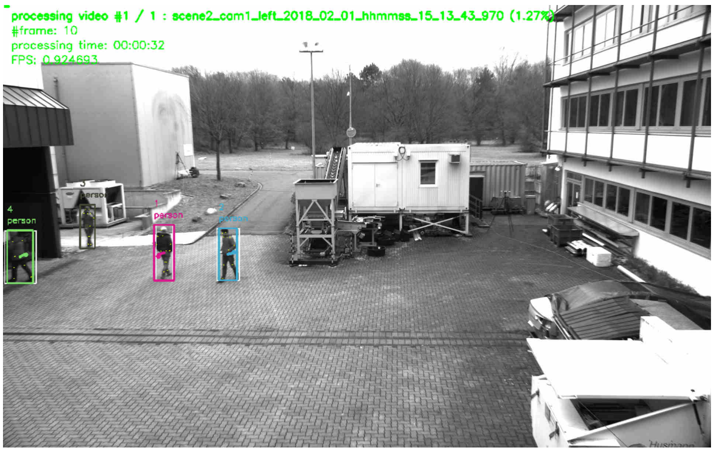

Figure 5.15: Frame taken from feature detection preprocessing procedure. During processing, the original video is played back while detected objects are highlighted.

As a first step in the feature extraction pipeline, each video from the input pool is processed in an object detection module (convolutional neural network (CNN)) that extracts all detected entities (persons and objects) for each frame of the video. In the present case, the YOLO v3 library (`https://pjreddie.com/darknet/yolo/`) is applied. The network was used pretrained on the mscoco dataset [226]. The result is a set of independent *Detections* for each frame, each containing information about its location in the image space (bounding box), a confidence score, and a classification of the object type (e.g., car, person, and backpack). In a second feature extraction step, each detection classified as "person" is processed in a skeleton extraction module by using the mentioned bounding box coordinates of the detection as input parameter. The current status of the proposed framework includes the OpenPose skeleton extraction library (`https://github.com/CMU-Perceptual-Computing-Lab/openpose`). Similar to the YOLO module, the OpenPose network was used pretrained on the mscoco dataset [226]. With that, each detection of the class "person" is enriched with skeleton information representing the key points of the detected skeleton. Figure 5.16 depicts a skeleton (right) as shown later in the exploration framework for the detection of a person (left). Subsequently, for each video, all detections are processed in a re-identification module, comparing the detections from different frames and identifying all detections that belong to the same entity (*Track*). The current implementation exploits DeepSort (`https://github.com/Qidian213/deep_sort_yolov3`), which reuses features from the object detection module (YOLO v3). As a result, a set of tracks is available for each video, each of which containing detections from the same entity (e.g., person and car) that describes its spatial movement over time. For each track, a representative detection is selected in which the entity is optimally represented (*Best Shot*). Currently, the detection with the highest confidence (YOLO output) is selected as best shot. Last, a global re-identification module is used to find relations between tracks of different videos. The detection features from the object detection module are used to compare the tracks of different videos and form sets of tracks that belong to the same entity.

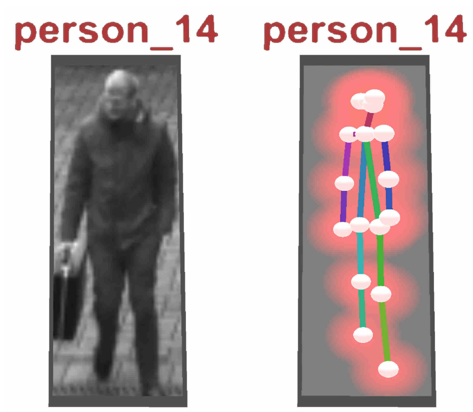

Figure 5.16: For each person recognized in a video frame (left), OpenPose is applied for skeleton extraction. The extracted skeletons can later be displayed in the scene as connected points (right).

Position Mapping to World Space After each video is passed through the object detection pipeline, the position information of the detections remains in camera space (frame coordinates in the pixel space). To locate the detection positions in the 3D scene, the corresponding 3D coordinates are computed based on the extrinsic and intrinsic parameters of the respective camera. These parameters are available due to the preceding scene reconstruction, pose estimation process, and camera characteristics. Depending on the information available, there are two different approaches for calculating the 3D coordinates.

The first strategy requires the availability of depth images containing the depth for each pixel of the respective camera frames. If this depth information is available, the detections' 3D position can be retrieved based on their bounding box position in the respective frames. This required depth information is available natively for binocular cameras that can capture 3D images [163]. For mono cameras, the depth information can be estimated as described in section 5.2.3.3. This means that in our case, depth maps exist for all videos that were processed in the dynamic scene reconstruction module.

As depth maps may be noisy and thus lead to faulty 3D localizations of detections, we provide an alternative strategy based on the 3D mesh created in the scene reconstruction step (see section 5.2.3.2). *Raycasting* is applied to calculate the 3D position of a detection. As shown in Figure 5.17, originating from the estimated camera coordinates (section 5.2.3.2), a ray (red line) is emitted through the bottom center of the detection in the image. The intersection point of the line with the geo-registered mesh is then used as the 3D position of the detection. To determine the angle of the ray, the estimated camera intrinsic properties (focal length and lens distortion) are used in combination with the camera space pixel coordinates of the detection. This approach is based on the assumption that objects must reside at the ground. With objects that are not at the ground, this strategy is, of course, prone to errors. For instance, if a person jumps, then this approach would calculate a wrong position for the time span during the jump, which is farther away from the camera than it is actually the case. To mitigate such inaccuracies in the future, it might be helpful to consider the position and direction of object shadows [392]. Unfortunately, such shadow algorithms depend heavily on lighting conditions. The consideration of shadow effects will therefore probably not eliminate all existing problems of this position extraction task.

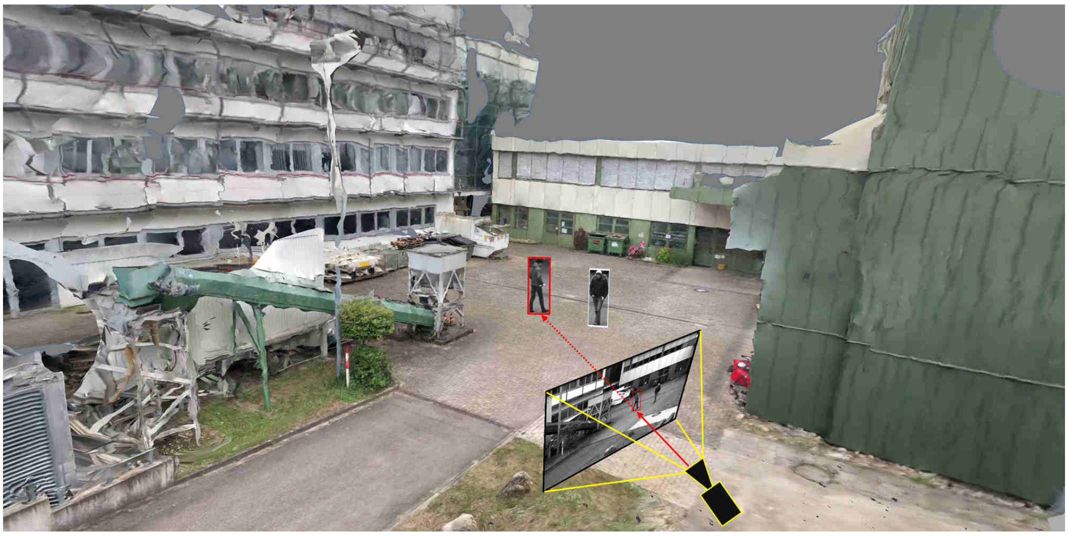

Figure 5.17: Its extrinsic parameters define the world coordinates of a camera in a 3D scene (camera icon). Based on intrinsic parameters, the pixel coordinate position of an object can be transformed into its respective world position through raycasting. A ray (red line) is emitted through the image at the lower edge of the bounding box of a detection (red rectangle in the camera frame). The intersection of the ray with the mesh provides the related 3D world coordinate.

Another promising approach builds on the availability of multiple cameras which capture the same detection from different angles. It would be possible to perform a triangulation of emitted rays from several cameras to detect the position of an object in the 3D world space [325]. Overall, the simultaneous application of different positioning strategies allows for modularity by prioritizing more accurate procedures. In the future, this modularity allows to extend the available methods, for example, by camera triangulation and shadow position estimation approaches. Eventually, the proposed prototype enables analysts to view the original footage, which is crucial for confirmatory analysis and critical decision-making processes.

5.2.3.5 Temporal Footage Synchronization

Another factor why a detailed analysis of video footage is time-consuming and costly is the inaccurate temporal synchronization of several cameras that usually originate from heterogeneous sources. Available video footage must be temporally synchronized so that analysts can get an overview of an incident. This temporal synchronization accuracy also affects the resulting quality of feature extraction methods that require multiple cameras (see section 5.2.3.4). The available footage is often not temporally appropriately synchronized, resulting in poor analysis results, which may even lead to false assumptions that impede accurate decision-making. Therefore, it is crucial to identify the correct temporal synchronization before the information is used in further analysis steps. Even minimal time differences may have a significant impact on critical decision-making processes.

The most basic strategy for performing time synchronization is to use the meta-data of a video to determine its start time. However, the information stored in video files is often incorrect. The system time in cameras may be inaccurate due to manual settings, or the stored creation time is overwritten when the video files are converted or copied. For small adjustments, analysts can manually manipulate the time offsets by to use appropriate values and adjust these offsets for each camera separately. However, this method is tedious, time-consuming, and error-prone, especially as the amount of the video material increases. Therefore, it is advisable to use available auditory or visual features that appear in the video content. For example, analyzing the cameras' soundtracks to extract distinctive audio features such as shots, shouts, or other significant noises could enable (semi-)automatic temporal synchronization to reduce the manual effort of the analyst. Additionally, it may be helpful to consider the visual features of video frames to identify similarities of events and synchronize the video material based on such anomaly conditions. For instance, the appearance of outstanding visual elements, such as a red bus driving through the scene at a specific time, could be used to match different videos temporally. Moreover, the trajectories of detected objects or persons could be compared and used for temporal synchronization. Enabling analysts to inspect the automatic temporal synchronization is essential for verification. For example, providing a time-aligned list of all videos enables the user to see and compare aligned frames.

So far, the proposed system only supports the first hands-on approach, which is feasible for scenarios with few cameras. As discussed, this approach does not scale for large numbers of videos, which would require the implementation of automatic algorithms that can be monitored by analysts. However, in line with the modular approach of the entire framework, we plan to integrate additional temporal synchronization modules, which can be selected in the preprocessing step depending on the available information.

5.2.3.6 Preprocessing Run Times

Preprocessing times vary and are highly dependent on the current constellation of modules, their configuration, the input data, and the underlying hardware infrastructure. The proposed pipeline comprises standalone modules that were benchmarked individually by their respective authors. Nevertheless, in the following, we provide a rough overview of preprocessing times for the previously presented constellation of modules on a conventional consumer desktop PC (GeForce GTX 1080Ti, 32 GB RAM, SSD, Intel i7-6700K). The considered data set consists of one handheld camera video (2 min, 1080 p) and three static camera videos from surveillance cameras recording a scene (1.5 min each, 1080 p). For the static 3D scene reconstruction, mainly frames from the moving camera are taken into account (two-minute video). A sparse reconstruction can be created within 10 min. The following dense reconstruction requires approximately 90 min to complete. It is noteworthy that this step only has to be completed once and is not affected by additional static cameras that are embedded in the scene. The given example data set comprises three time-dependent videos which are reconstructed as a dynamic scene. The dynamic object detection and segmentation, as well as the monocular depth estimation, requires about 250ms per frame. For the given example data set, this results in about 27 min for 6480 frames. This process can be sped up by only reconstructing keyframes. In particular preprocessing times of the high-level analysis are highly dependent on the content of videos, i.e., if many objects appear in the scene, the time increases, and vice versa. Object detection, re-identification, skeleton extraction, and 3D pose calculation require roughly 500ms per frame. For the given example dataset, this results in an overall run time of 54 min. In the current configuration, the modules were opted for high-quality results. By tweaking parameters, for instance, by disabling multi-resolution object identification in the YOLO module, processing times can be sped up significantly.

In summary, the preprocessing pipeline required approximately 181 min to process the example data set and display the result in an enriched 4D scene, which can be explored interactively. To allow for fast analysis procedures, all individual modules can be tweaked at the price of lowering the quality of results. Additionally, thanks to the modularity of the pipeline and highly parallelizable modules, preprocessing computation can be outsourced to more powerful GPU clusters, shrinking preprocessing times to a fraction of the ones described above.

5.2.4 Visual Exploration of 4D Reconstruction

After completing the preprocessing pipeline, analysts can inspect and examine the reconstructed scene using an interactive application. Figure 5.18 depicts the main building blocks of the analysis application. On the left is the 3D reconstruction, including a static mesh of the given environment and spatially registered time-independent materials such as photos and panoramas. This environment can be spatially explored and enriched with annotations, even if no time-dependent materials were added to the analysis. On the right are all time-dependent materials, such as cameras with estimated locations in space per timestamp, dynamic point clouds, and extracted meta information. We provide a video to demonstrate an exemplary visual exploration of a 4D reconstruction (`https://www.youtube.com/watch?v=bcDrLCaI2RI`). In the following, most examples are taken from a 4D reconstruction based on the dataset provided by Pollok [288]. The dataset includes several scenes in which several persons, cars, suitcases, etc. are visible and actors reenact different scenarios (e.g., kidnapping and dropping suitcases). It contains video material from three static surveillance cameras as well as footage from handheld devices monitoring the reenacted incident.

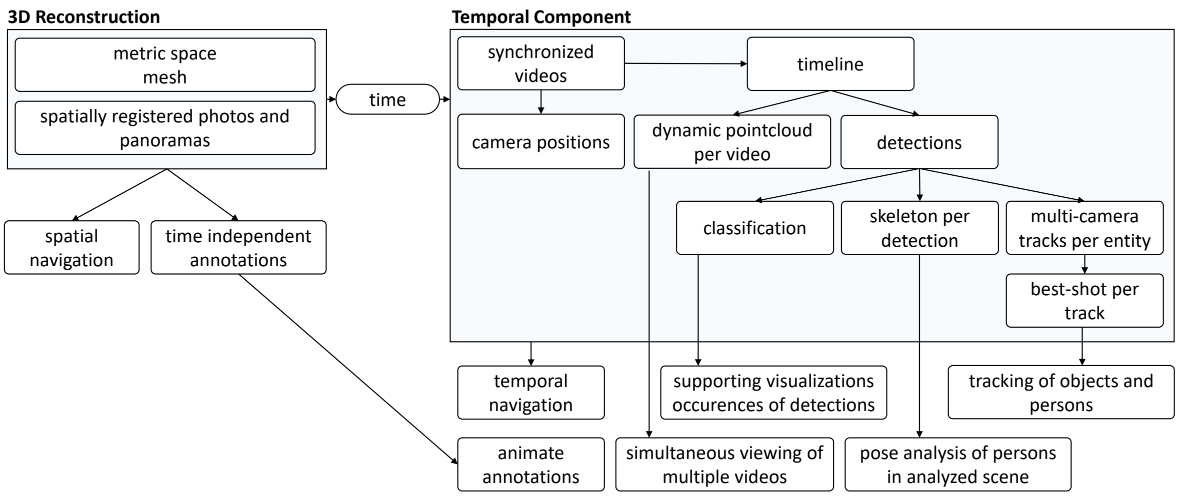

Figure 5.18: Main elements of the analysis application.

Figure 5.19 gives a first impression of the benefit of the presented approach. The static 3D reconstruction serves as a base visualization in which all input sources can be placed within a shared context. Static elements such as photos (green) and panoramic images (teal), which provide additional contextual information about the environment, can be spatially registered. Their positions are indicated as camera icons or spheres. Video sources are also spatially registered and visualized as camera icons (orange, blue, and red). This visualization gives the user a good overview of all available input sources and their spatial distribution. It allows the user to relate sources to each other and, for example, find all cameras directed to a certain point of interest. Last but not least, the video footage is temporally synchronized and automatically extracted meta information from all sources can be displayed simultaneously. Blue and yellow dashed lines mark skeletons of detected persons, which were found in different videos. In the following, the building blocks of the demonstrator are described in more detail.

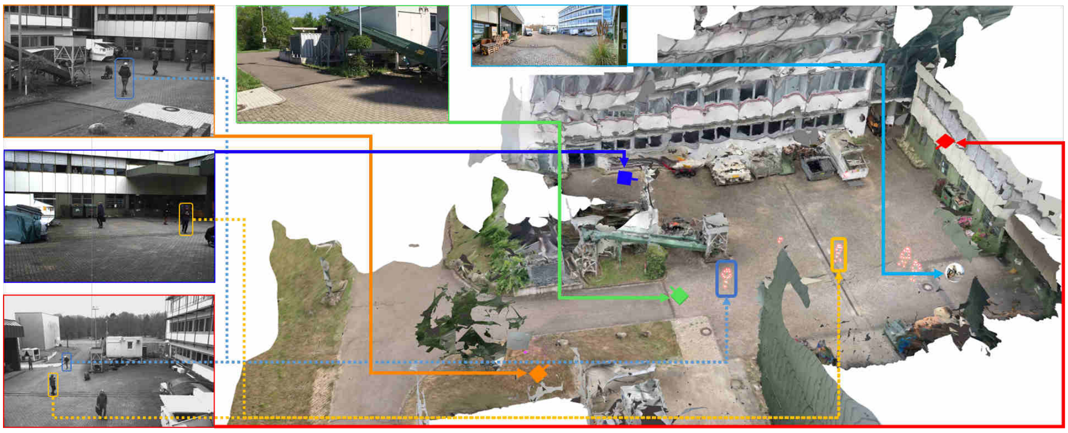

Figure 5.19: Multiple data sources are bundled and displayed simultaneously in a shared context. On the left side, three frames from static surveillance cameras are displayed. Their locations are indicated by small camera icons in the 3D scene (orange, blue, and red). Detections from all cameras are displayed simultaneously in the scene (dashed lines) as well as static material, such as photos (light green) and panoramic images (teal).

5.2.4.1 GUI

The ISRA-4D interface for the visual exploration of the 4D reconstruction comprises four main parts. As shown in Figure 5.20, the view of the reconstruction (center) is surrounded by three panels: a menu bar at the top, a mini-map in the top right corner, and a timeline panel at the bottom.

Menu Bar The menu bar at the top (see Figure 5.20) allows the user to configure the appearance of the inspected scene and provides options for additional user interactions. A time slider at the top left corner with a play/pause button is followed five menu panels. The first panel (*Layer Options*) allows the user to switch visual layers, such as static and dynamic point clouds, the static mesh, or camera icons in the main scene. The second tab (*Custom Annotations*) provides functions for adding, loading, saving, and changing manual annotations. The third tab (*Real Data*) contains three interaction options for entering or retrieving original photo/video footage into or from the scene. Next to it, there is a panel (*Detections*) for configuring the appearance of automatically extracted content, i.e., detected persons and objects. For instance, users can determine whether bounding boxes should be displayed or skeletons should be drawn into the scene. The last tab (*Advanced*) contains additional functionalities for manually editing and saving automatically extracted detections as well as options for configuring the VR interface.

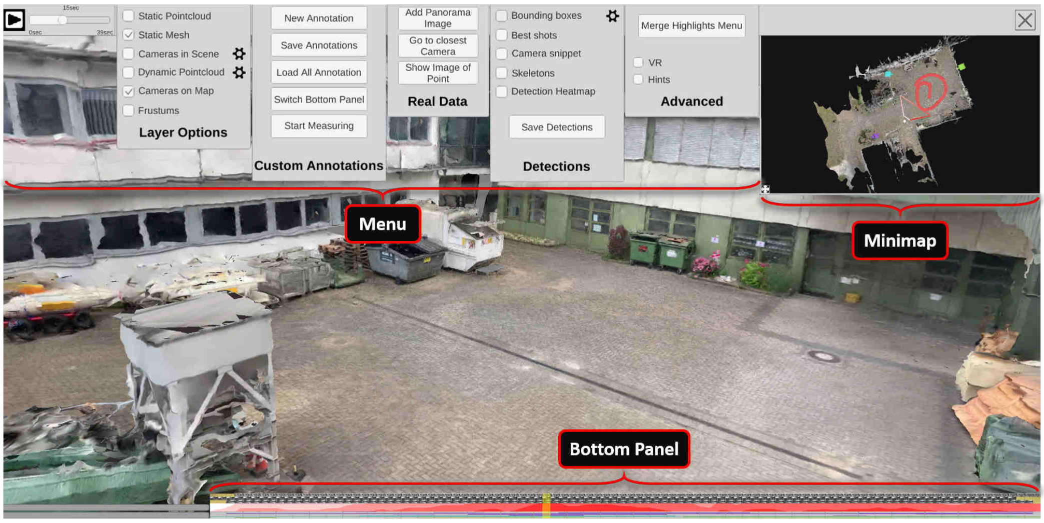

Figure 5.20: The graphical user interface of the presented demonstrator consists of four main parts: a menu at the top, a minimap at the top right, a bottom panel, and the main window as a view of the inspected scene.

Minimap To keep an overview while inspecting the scene, a minimap at the top right provides a birds eye view of the environment (see Figure 5.21). The current position and viewing direction of the observer is indicated by a red dot and a frustum of pyramid. The minimap's size can be arbitrarily changed by using drag and drop on the small icon at the bottom left. If cameras are active at the currently selected time, their icons can also be displayed in the minimap, giving the observer an overview of all sources that were monitoring the scene at the selected time. Selecting a camera in the minimap changes the viewport of the main window to that of the selected camera, allowing the user to inspect the scene from the perspective of the source and, if desired, view the original video footage.

Figure 5.21: Minimap depicting a top-down view of the reconstructed environment. The locations of the cameras recording the investigated incident are displayed as small camera icons (3 static cameras: blue, green, and magenta; 2 moving cameras: red and yellow). The current location of the user is shown as a small dot, with a red frustum indicating the viewing direction (center) and field of view.

Bottom Panel The bottom panel represents a timeline that covers the period from all time-dependent contents (see Figure 5.20, bottom). By default, it is collapsed, but it unfolds when the mouse cursor is moved over it, revealing frame previews and additional visualizations about the class distribution and the duration of detections in the scene. See section 5.2.4.4 for more information on the visual elements in the bottom panel. A transparent yellow slider indicates the current temporal position. By clicking on the timeline, the analyst can choose to view a specific point in time manually.

5.2.4.2 Reconstruction (3D) & Spatial Navigation

Time-independent elements comprise a static 3D reconstruction of the environment, including photos and panoramas that are temporally not registered, and manual annotations (see Figure 5.18, left). All these elements are static and can be explored independently of temporal navigation. The static 3D reconstruction of the environment serves as a base visualization. Figure 5.22 shows an exemplary 3D reconstruction. The user can navigate through the scene using standard input modalities (mouse and keyboard). The virtual camera can be moved with the keyboard and rotated with the mouse.

Figure 5.22: 3D scene that can be inspected by flying around in it, which interactively changes the perspective.

Photospheres and Time-Independent Materials Additional materials collected for the respective environment can be embedded into the scene, including forensic evidence photos of a crime scene, panoramic shots, and photospheres. These expand the context provided in the analysis, or the analyst can employ them for pre-post comparisons. Figure 5.23 shows how a panoramic image (left) is displayed within the 3D scene as a textured sphere (center). Once the user clicks on the sphere, the virtual camera is moved to the location of the sphere and the panorama is blended over the 3D reconstruction (right).

Figure 5.23: Panoramas (left) are displayed as spheres in the scene (center). By opening a sphere, the user “enters” the photosphere to inspect it (right).

The user can then “look around” in the 3D scene by similarly using the mouse as before. Photos, photospheres, and panoramas can also be inserted live during the entire analysis process. For example, if the automatic positioning of panoramas is not correct or if new sources became available after the reconstruction was completed. For that to happen, the user can press the corresponding button in the top menu (see Figure 5.20, “Real Data”), position the new sphere, and select the image from the hard drive to be inserted.

Annotations Users can further enrich the scene by manually adding static annotations (see Figure 5.24). The respective interaction for adding a new annotation can be selected in the top menu bar. The user then defines the position of the annotation in two steps. First, a horizontal plane must be selected, which determines the height of the annotation. Second, users can select a position on the plane, allowing them to pick the desired 3D location on the 2D screen. After setting the location, width, height, and depth, additional information of the annotation can be altered. For example, the user can specify a custom logo, enter an annotation label, and type notes.

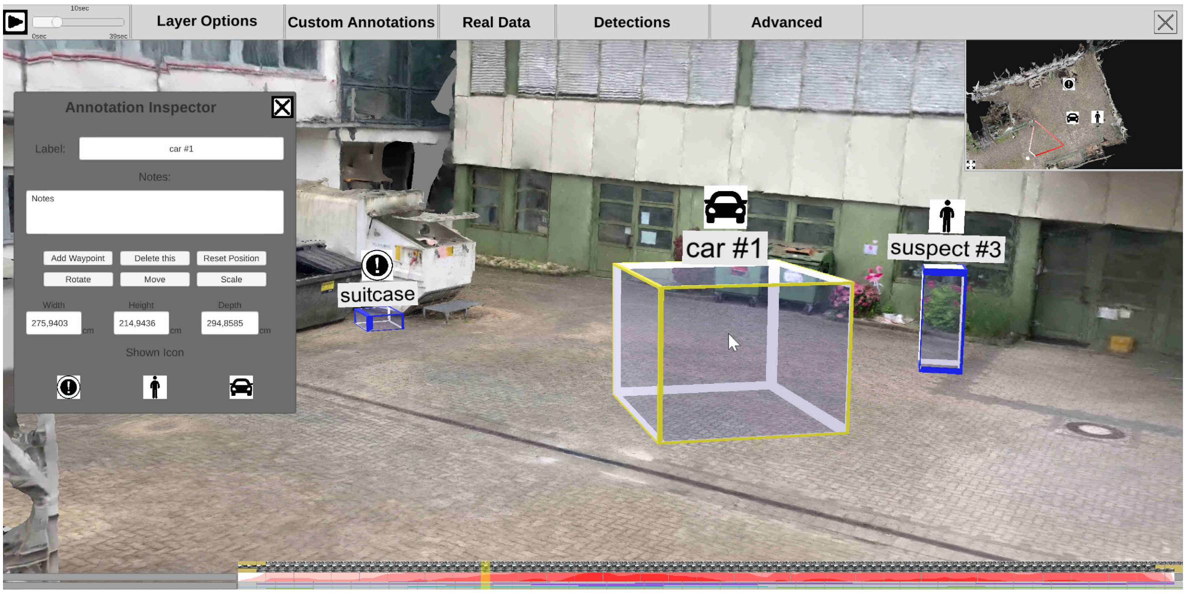

Figure 5.24: Static user annotations can be manually added to the scene.

5.2.4.3 Dynamic Content

Time-dependent elements include temporally registered cameras (static and moving), dynamic point clouds, detections from video footage, and animated user annotations (see Figure 5.18, right). All these elements are registered on the global timeline and can be explored by temporal navigation.

Temporal Navigation & Timeline For easy access, the timeline is displayed twice—at the top left corner and, in large, in the bottom panel (see Figure 5.20). The currently selected time is indicated by slider bars in the respective timelines and numbers in the upper left-hand corner. The time range is automatically extracted from time-dependent sources as the time span from the global minimum time to the global maximum time. To navigate through time, the user can drag the timeline handles or click anywhere on the timeline. With the play button in the upper left corner, the time can be automatically increased continuously, analogous to the real-time progression. Once clicked, the button turns into a pause button that can be used to stop the automatic increment in time.

Camera Positions All time-dependent video sources are registered in the timeline, and their location is indicated by a small camera icon that can be made visible if required in the scene and the minimap, separately. Camera icons are only visible if they provide footage for the currently selected time on the time slider. As shown in Figure 5.25, different types of frustums can be displayed, highlighting the area in the 3D reconstruction that is covered by the video material of the respective camera. On the left side, a semi-transparent frustum is inserted into the scene, clearly showing a cut through the scene where the viewport ends. On the right are two alternative frustum options that work with illumination. At the top right, the camera projects a colored light into the scene that illuminates everything seen by the camera. At the bottom right, a subtractive approach is shown, which hides everything that the camera cannot see.

Camera icons in the scene and on the minimap can be clicked to take a look at the scene from the perspective of the selected camera. Once a camera is selected, it is possible to view its original video material. Additionally, the user can select interactions from the menu bar option “Real Data” to jump to the nearest camera to discover original footage for a potentially interesting perspective. The user can also select a point on the 3D reconstruction to retrieve all cameras with the chosen point within their field of view at the currently selected time on the time slider.

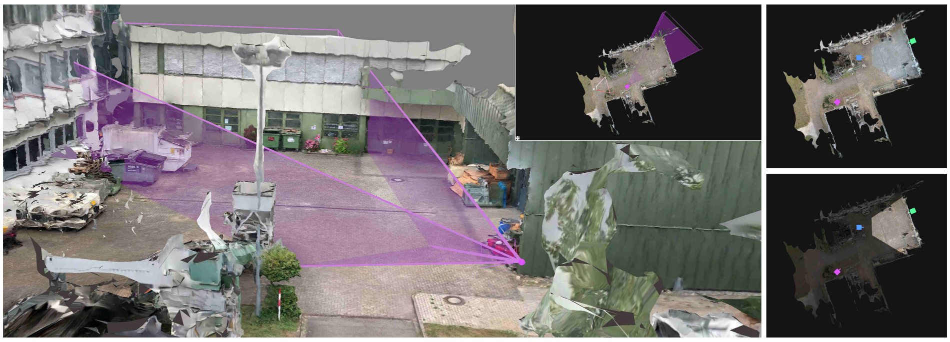

Figure 5.25: The camera frustums displayed in the scene and minimap can be customized: either as semi-transparent objects (left) or using additive (top right) or subtractive (bottom right) lighting.

Moving cameras change their location in the scene over time. Therefore, a camera icon is created for each frame, as shown in Figure 5.26. The camera position at the currently selected time is highlighted with a red halo around the respective camera icon. This view helps to get an overview of where a camera has moved to and which areas are generally covered. The user can configure the display of moving camera frustums to reduce clutter. The user can set for how long after and before the currently selected time dynamic frustums should be displayed. This makes it possible to show only one frustum that changes its location over time or to show the current one with any number of preceding and succeeding camera icons. Transparency is used to encode temporal distance.

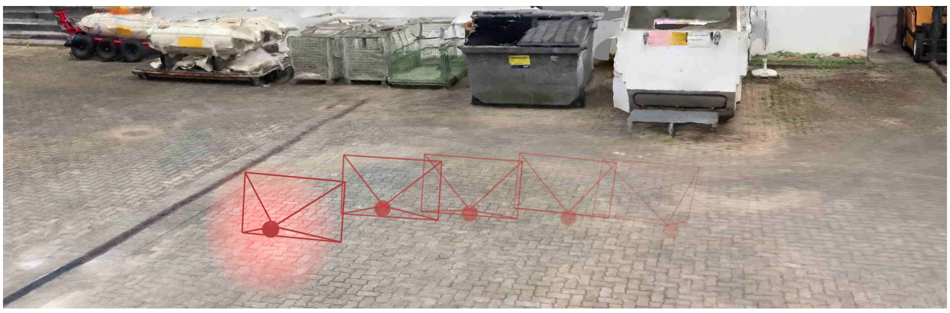

Figure 5.26: The user can configure the display of moving cameras in the scene. The location of the camera at the currently selected time is highlighted with a red halo. In this example, the camera locations of the last four time steps are also shown with increasing opacity.

Detections Each time-dependent video is preprocessed in a feature extraction pipeline (see Sectiohn 5.2.3.4). The generated information can be visualized in the 4D reconstruction. Using the top menu bar, the user can configure how detections are displayed. As shown in Figure 5.27, a detection can be represented as (a) an abstract minimum bounding box with a title; (b) the best shot of its track; (c) the snippet of the minimum bounding box from the original video, as (d) a combination of a, b, and c; or, in case the detection belongs to the class "person", (e) a skeleton. Detections are displayed with regard to the currently selected time on the time slider. This means that all detections found in different sources at the selected time are displayed. The user can select the sources (cameras) from which detections should be displayed. Additionally, it is possible to filter detections according to their class, so that only detections of certain classes, such as suitcases and bikes, are displayed.

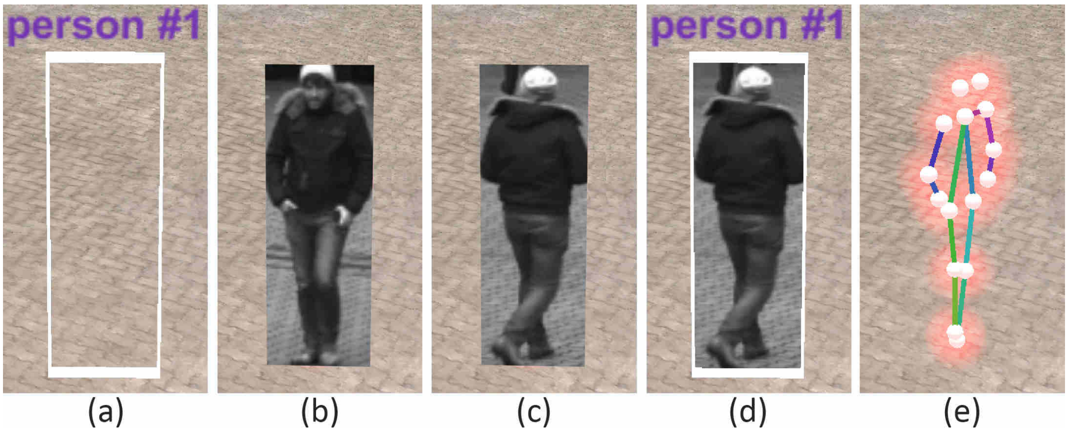

Figure 5.27: (a) A detection can be displayed as a bounding box, (b) the best shot of its track, (c) the corresponding snippet from its frame, (d) a combination of bounding box and best shot or frame snippet, or, if available, (e) its skeleton.

By clicking on a detection in the scene, a context menu appears on the left side (see Figure 5.28). This menu allows the user to change its label and add notes. Once a detection has been selected, its track trajectory is visualized in the 3D scene, depicting its spatial progression over time (see Figure 5.28). Black arrows on the trajectory indicate the direction of movement.

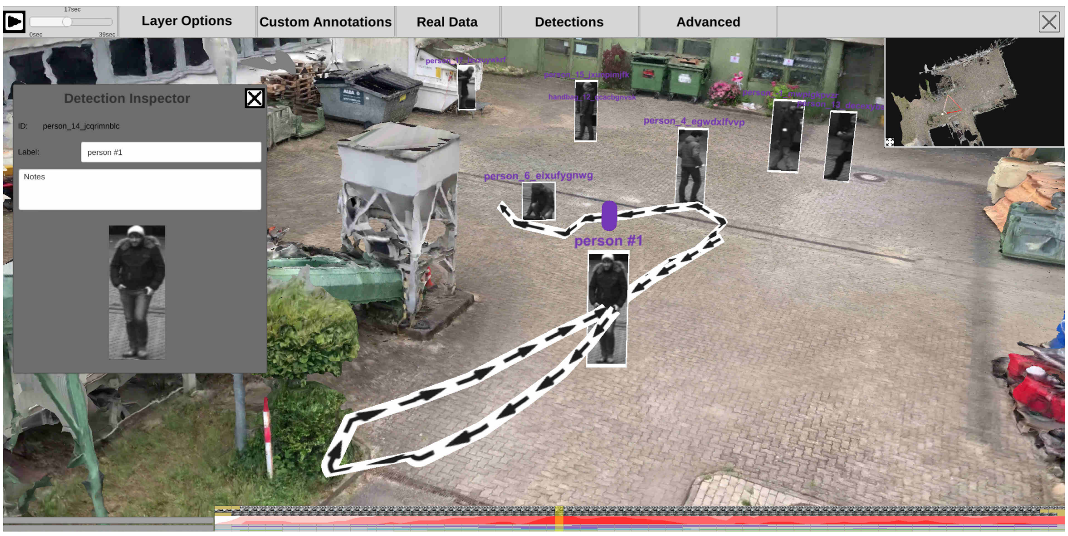

Figure 5.28: The trajectory of a selected detection is visualized as a directed path within the scene. A menu allows to change the displayed title of a detection and to leave notes.

The demonstrator comprises tools to manually refine and adjust automatically extracted information in case the automatic approach did not work as desired. If, for example, two persons who look similar cross in front of a camera, their tracks may get mixed up, and the information displayed is incorrect. The user can solve such issues by splitting the respective tracks and then merging the related tracks.

Identities can be anonymized to protect the privacy of persons in the 4D reconstruction. For instance, if the tool is deployed for an investigation at a public place, faces of bypassers who are not subject to the investigation itself can be pixelated, making them unrecognizable. As displayed in Figure 5.29, a face detection algorithm detects bounding boxes of faces, and the respective area is pixelated. Throughout the analysis, investigators can reveal the faces of persons that are relevant to the case. A right management system could be deployed to regulate face revelations and define who can use this function.

Figure 5.29: Faces of displayed persons within the 4D reconstruction can be anonymized for privacy reasons. A face detection algorithm detects the bounding boxes of faces (center) which are subsequently blurred in the displayed content throughout the visual analysis (right).

Dynamic Point Clouds For each time-dependent camera in the scene, a dynamic point cloud is extracted in the preprocessing step. During the inspection of the scene, the 3D point clouds of all cameras recorded at the currently selected time can be displayed simultaneously. In this way, it is possible to perceive the content of several videos at the same time and in a mutual spatial context without much mental effort. While playing, the observer can fly through the scene and observe the progression of events from different perspectives.

Of course, it is possible to select which cameras are to be displayed as 3D point clouds. The advantage of this technique is that, unlike automatically extracted meta-information (e.g., detections), the original video footage is completely mapped into the 3D scene. Due to detection or classification errors, certain people or objects may not be detected in a frame and, therefore, not be displayed in case the option to show all detections in the scene is selected. However, if the point cloud of the respective frame is visualized in the scene, each pixel of the input frame is also displayed.

If the point cloud is viewed from the camera location, it resembles the original video footage. Figure 5.30 shows an example of the dynamic point cloud visualization in the reconstructed environment. In the current perspective, point clouds are displayed as shown on the left. However, when navigating through space and observing, for instance, a person from different perspectives, different point clouds can be observed from cameras monitoring the scene from respective directions (e.g., (1) from the orange and (2) from the blue viewing angle).

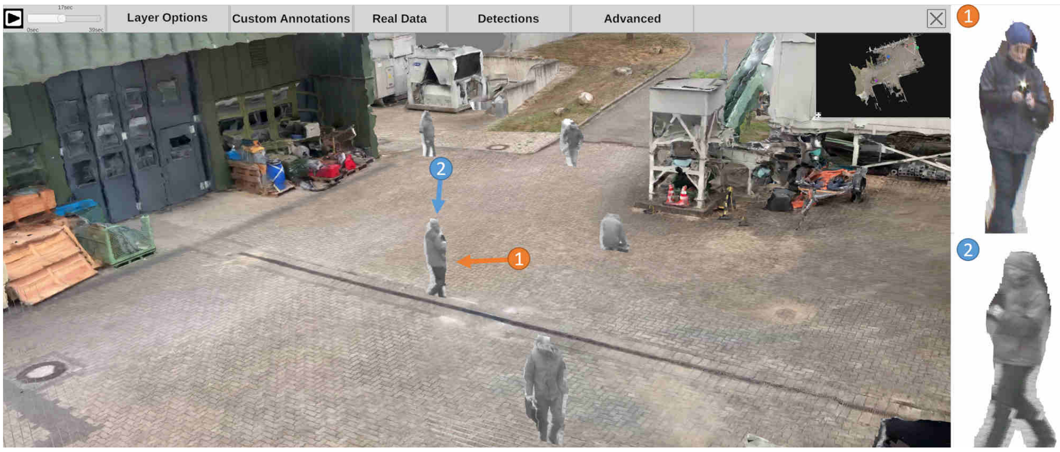

Figure 5.30: Dynamic point clouds displayed in the static scene from the current perspective (left). If one navigates through space, the perspective changes and point clouds generated from different cameras can be perceived. For example, (1) the (top right) point cloud snippet can be seen from the direction indicated by (1) the orange camera and (2) the (bottom right) one from the direction indicated by the blue camera.

Animated Annotations Static annotations can be animated and thus integrated into the global timeline along with videos and detections. To do this, the user can select the annotation to animate and select the option to add a waypoint. This action creates a waypoint for the annotation's current location, and another can be set interactively, similar to the location selection of annotations (see section 5.2.4.2). When adding waypoints, the bottom panel view is automatically changed to a waypoint timeline view (see Figure 5.31). Each annotation is displayed in a list, and all corresponding waypoints are lined up for each annotation. The user can shift waypoints in time to define the location of the annotation at a specific time. If two waypoints are at different locations, the annotation position is interpolated depending on the time between the two waypoints. In this way, the user can, for example, reconstruct a course of events described by eyewitnesses or plan the progression of an intervention.

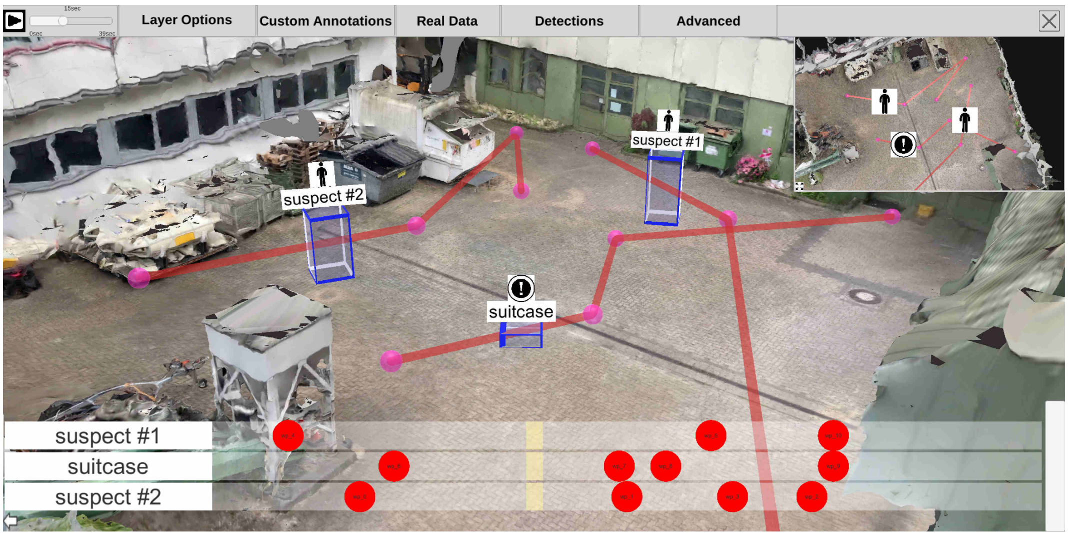

Figure 5.31: To animate annotations, waypoints can be set and arranged on a timeline that temporarily replaces the bottom panel. Waypoints determine the location of an annotation at a particular time.

5.2.4.4 Visual Analysis

The bottom panel can be expanded by hovering over it with the mouse, as depicted in Figure 5.32. Optionally, it can be pinned to stay open and reveal the underlying visualizations while navigating through the environment. When it is enlarged, three visualization elements appear. At the top, all time-dependent video frames are aligned in a scrollable and filterable list that indicates the start and endpoint of each video on the global timeline and provides a first glance at the original video footage. The visibility of the frames is supported by a fisheye effect induced by hovering over the frames. Dependent visualizations accordingly scale while hovering.

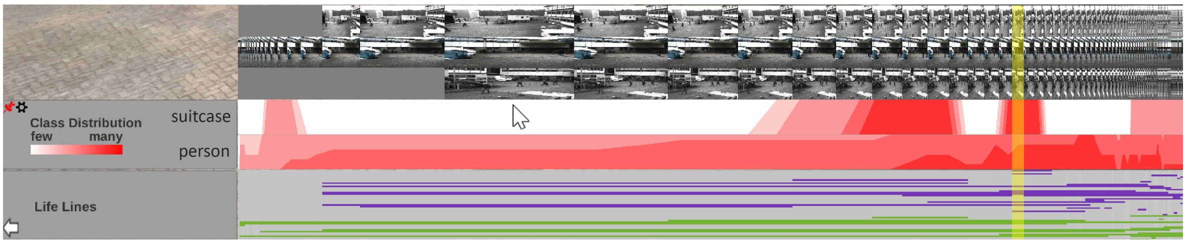

Figure 5.32: The bottom panel consists of three elements: At the top is a frame preview of all selected cameras. In the center, the class distributions of the detections are visualized as horizon charts. At the bottom is a chart depicting the appearances of all detections as lines.

The other two elements (center, bottom) are visualizations of automatically extracted detections. A horizon chart visualization shows the distribution of a particular class of detections over the entire time axis in the center. The user can select classes and combinations thereof to be displayed as individual horizon charts. For each selection, a horizon chart is displayed, showing the total number of detections of the selected class(es) at each time. In the example (see Figure 5.32), one horizon chart was created for all detections of the “suitcase” class and another for the “person” class. With the given visualization, it is easy to identify at what times a particular object type was detected to jump to it quickly. Besides, this visualization allows the user to identify periods when nothing was detected, which in surveillance use cases, for example, helps to sight large amounts of video material more quickly.

The last visualization element (bottom) shows the time spans of detected tracks. Each row belongs to a single object and is filled with color when it was detected in the original footage. The line’s color represents a visual link to the corresponding video in which the track was detected (colored camera icons in scene and minimap). The sources of the detections visualized in the bottom panel can be filtered as desired to analyze only one video or a subset of videos.

Additionally, the current selection of detections (after filtering by classes and video sources) can be displayed as a heatmap projected onto the reconstruction (see Figure 5.33). Once the user selects the respective option in the top menu bar (Figure 5.20, “Detections”), a 2D heatmap with the locations of all selected detections is created in the background. A grid of 10×10 cm tiles with the environment’s size is used to create the heatmap. Each tile counts how many detections are in its corresponding area. The resulting $n \times m$ grid is smoothed with a Gaussian kernel, normalized, and finally saved as an image in which the corresponding tile value determines each pixel on a user-defined color gradient (from transparent to yellow to red). The generated image is then projected orthogonally onto the mesh.

With the heatmap visualization, the user can quickly identify which areas of the environment contain most occurrences of persons and objects and which areas were uneventful. For example, by filtering for all detections of the class “suitcase” and displaying the heatmap for it, the user can quickly see where a suitcase was detected to find all video sources which monitored the respective areas. Especially in combination with horizon chart visualizations, it is easy to find all videos in which and the corresponding times at which suitcases were spotted.

Figure 5.33: A heatmap visualization of selected detections can be projected onto the environment providing an overview of where objects or persons were detected in the analyzed scene.

Throughout the analysis, the user can insert images or spatial and time-dependent annotations. Additionally, the user can activate a measuring tool in the top menu (Figure 5.20, "Measure Distance"). Once activated, measuring points can be created by clicking on the environment (see Figure 5.34). Intermediate segments are labeled with their distance in meters, and the total length is displayed in the top menu.

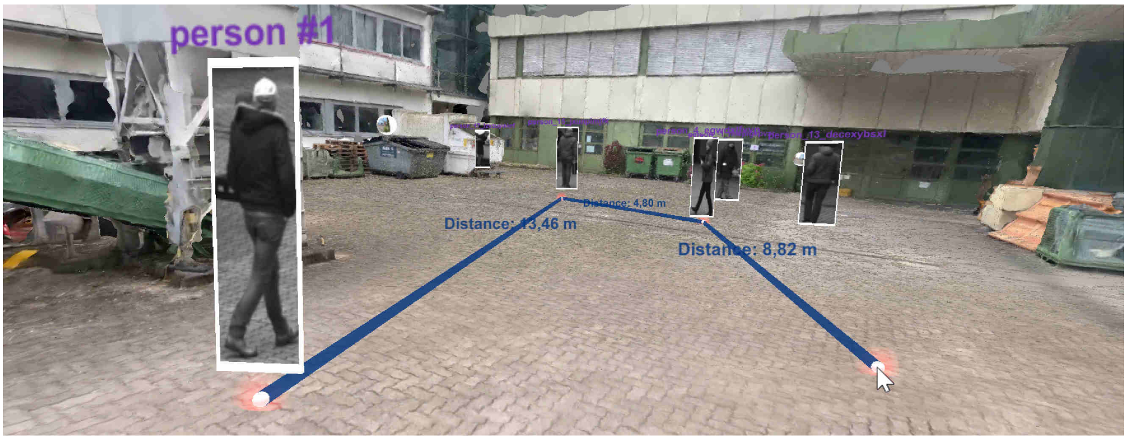

Figure 5.34: Interactive tool for measuring distances and object sizes in the reconstruction.

5.2.4.5 VR Exploration

Besides exploring the 4D reconstruction on a monitor screen, it is possible to enter the scene in VR. When a VR headset is connected to the PC, it is automatically recognized and configured for usage. In the current example, we use a Valve Index VR HMD [373]. The environment is scaled to metric space, which means that the environment is represented as a life-size model. Distances and dimensions of objects can be viewed as in the real world. Figure 5.35 shows an example scene as it can be observed in VR. All visualization elements like detections, heatmaps, camera positions, and annotations can also be inspected in VR. The user can navigate through space by walking (if the available physical space allows it) or virtual teleportation. The user can press and hold the touchpad on the right controller to select a target location to teleport. When released, the user's location is set to the respective position.

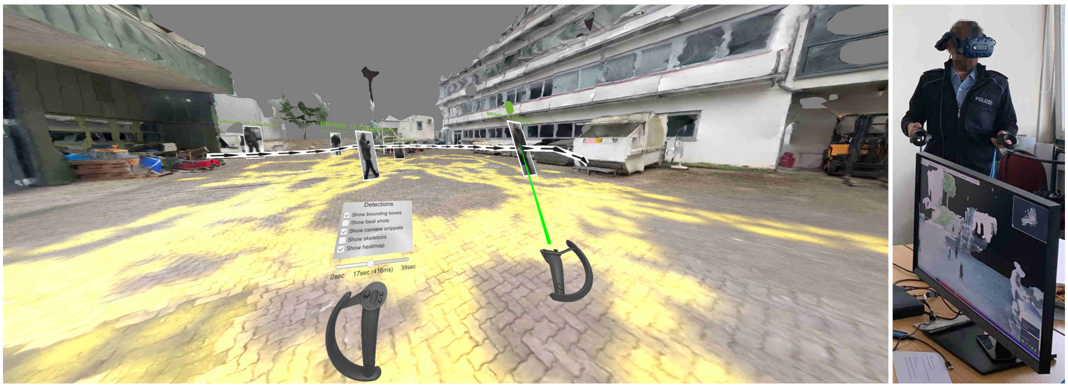

Figure 5.35: Left: View of an exemplary scene in VR. Right: Set-up with immersed investigator.

For further interactions, the user can open a menu by placing the thumb on the right joystick (see Figure 5.36, left). The joystick can be moved around and released at the desired option to select an item on the radial menu. The selected menu opens and is attached to the left controller. Options can then be selected by pointing the laser on the right controller at them and pressing the right trigger button. In this way, it is possible to modify the appearance of the scene ("Layer Options"), the display of detections ("Detections"), and the annotation of the scene ("Annotation Options"). The user can either drag the slider attached to the left controller or use the left joystick to scroll forward and backward to navigate in time.

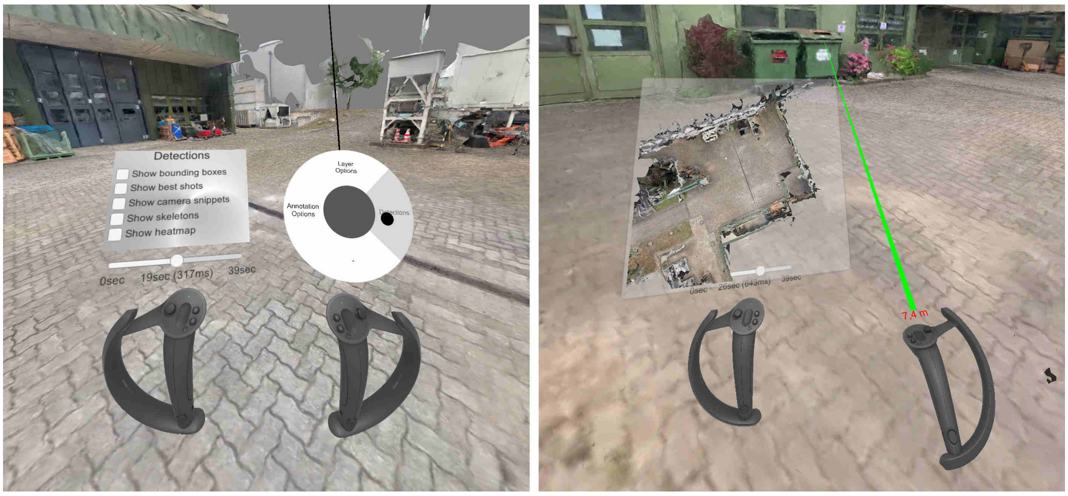

Figure 5.36: Left: A radial menu can be opened on the right controller to open various menus that are displayed on the left controller to configure the visualized scene. Right: A minimap and a distance measuring tool can be activated on demand.

A minimap of the environment can be toggled using the "B" button on the left controller (Figure 5.36, right). The user can click in the minimap to teleport to the selected location. Furthermore, the "B" button on the right controller can be used to toggle a measuring tool. After activation, the laser emitted from the right controller measures the distance to the surface that it hits and displays it above the controller.

Similar to interaction possibilities on the screen, the user can select detections and annotations in VR by pointing at them and pressing the trigger button. When selected, the respective menu is displayed and attached to the left controller. As traditional text input in VR is quite cumbersome, a speech-to-text module allows to change labels or add textual notes to annotations and detections via voice commands. For this, the user can select an input field, hold down the left trigger button, and record spoken input. Afterward, it is converted to text and inserted into the chosen input field.

Our exploration tool allows the simultaneous usage on a screen and in VR. While one collaborator observes the scene in VR, the other can interact as usual on the monitor while seeing the VR user as an avatar walking through the scene (see Figure 5.37, left). If two monitors are connected to the PC, one depicts the scene's regular interface and the other the observer's view in VR (Figure 5.37). The VR user can activate a laser pointer on the right controller by pressing the right trigger button to point at something for improved communication.

Figure 5.37: Collaborative setup with multiple monitors connected to the system. One monitor shows the usual view of the 4D reconstruction (left) and the other one, a view from the simultaneous observer's perspective in VR (right). An avatar of the VR observer is displayed in the desktop interface (left).

In a initial qualitative assessment, law enforcement officers who had the opportunity to test our demonstrator provided feedback (results of the interviews are reported in more detail in section 3.2.5.5). Eleven criminal investigators of the German Federal Police (Bundespolizei) evaluated the presented demonstrator. Overall, they were convinced of the added value of virtual reality in the given context. The main argument was that it was advantageous to inspect the scene "from within", as one could perceive the environment in its natural size and explore it by walking around in it. This made it easier for them to estimate distances between detected entities and get a better spatial sense of the scenario. Another potential benefit mentioned was that such virtual tours in VR could be used to inspect crime scenes from a distance without visiting their actual physical location, or to use them at court to illustrate a sequence of events in a criminal incident graphically.

5.2.5 Use Cases

In the following, four use cases demonstrate the potential of the proposed approach.

5.2.5.1 Mass Data Analysis & Preparation of Evidence

After major criminal incidents, such as terrorist attacks in a city, law enforcement agencies collect large amounts of evidence. These usually consist, among other things, of surveillance videos and recordings of eyewitnesses (photos and videos). In the case of unusually large incidents, law enforcement agencies even tend to set up platforms for the civil population to upload witness photos and video recordings. In total, the amount of digital information collected can easily exceed thousands of gigabytes of data and months of non-stop video recording. In practice, the review of evidence materials is still primarily done by hand. Criminal investigators go through all photos and videos, assess their relevance, and try to relate them to other materials, e.g., by annotating time, location, and meta information.

However, to cope with such an amount of data and facilitate the preparation of evidence for use at court, automatic mechanisms could be used. Our demonstrator could filter spatially and temporally relevant sources, place them in a shared context, and extract valuable meta information. To achieve this, the region of interest would first have to be specified and reconstructed as a 3D model. To return to the example of a terrorist attack in a city, videos of all affected areas could be employed to create a reconstruction. This can either be videos and photos from the time after the incident (filming the area with handheld cameras or drones) or video footage from the incident, which can be assigned with certainty to the area of interest (e.g., surveillance cameras with known locations). Subsequently, the entire pool of videos and images of the incident is fed into the reconstruction pipeline, trying to find matching points with the original reconstruction. Materials that do not overlap with the reconstruction are filtered out and may be subject to manual inspection. However, for all videos and images that can be registered at the site of the event, their location can be estimated and positioned in the reconstruction. After a semiautomatic temporal synchronization step, footage from before or after the incident can also be filtered out. All remaining evidence can be aligned on a shared time axis.

The 4D reconstruction can then be inspected, providing an overview of the environment, all available video and photo sources and their locations, and the time during which each location was monitored. This approach could drastically reduce the materials that need to be viewed manually. In addition, the approach facilitates the process of bringing recordings of the same area from different angles into a mutual spatial context without much mental effort. Moreover, the automatic approach could assist investigators in the analysis of meta information. For example, it would be possible to track an object or person through a single video or complete footage. In the graphical exploration, an entity's spatial progression would be displayed as a single continuous path, regardless of the recording's source.

Besides providing a quick first overview of the evidence, it can be used as a starting point for further adjustments, such as the manual insertion of footage that could not automatically be registered correctly. When all relevant information is in one place and can be located precisely in space and time, it is much easier to keep track of large amounts of evidence.

5.2.5.2 Crime Scene Investigation

Crime scenes, such as a murder scene, are carefully documented during criminal investigations. After the incident, the police collects forensic footage to record evidence and store as much information as possible about what the location looked like shortly after the crime was discovered. If available, information at the time of the incident is also considered, such as nearby surveillance cameras, witness reports, or even videos and witnesses' photos.

A 3D model of the crime scene can be reconstructed with the presented demonstrator, showing the crime scene as it was found before clean-up. All available sources are registered in the reconstruction. In this way, the highly unstructured mass of digital information is spatially organized, so that, for example, all sources recording a certain point in space can be easily identified. The 3D model could help criminal investigators organize the available information and put it into a spatial context without much mental effort. It also allows remote inspecting of the crime scene without physical presence and at a later time. For example, as pointed out by interviewed criminal investigators, the reconstruction could be used as an interactive, graphic basis for conveying information at court hearings.

Besides providing a permanent image of the crime scene, footage of the incident itself, if available, can be interlaced. Similar to the previous use case, footage from surveillance cameras and eyewitnesses can be displayed in the reconstruction. Meta information such as person and object detections and their tracks could be automatically extracted and displayed. As the forensic material provides much information about the scene after the incident, pre–post comparisons with the footage could easily be made. For example, suppose a specific point of interest was recorded with a surveillance camera. In that case, this point can be selected in the 3D scene to reveal all original footage sources which contain the same location.

The 3D reconstruction could also be incredibly helpful in reconstructing the sequence of events that led to the crime and a possible later course of events. If, for example, witness reports are available, criminal investigators can resort to animated annotations and try to visually model the described occurrences—possibly together with the witnesses themselves. Having a graphic representation of the environment in front of them could help them remember the course of events more accurately.

5.2.5.3 Real-Time Surveillance Scenario

The approach presented could also be used for real-time surveillance tasks. A monitored complex, such as an airport site, can be reconstructed as a 3D model—either with the previously presented method using large amounts of images or with alternative approaches, such as native 3D modeling (e.g., with a floor plan of the building), or 3D laser scanning. All available surveillance cameras are then spatially registered within the model, and their video streams are fed into the system (see Figure 5.38).

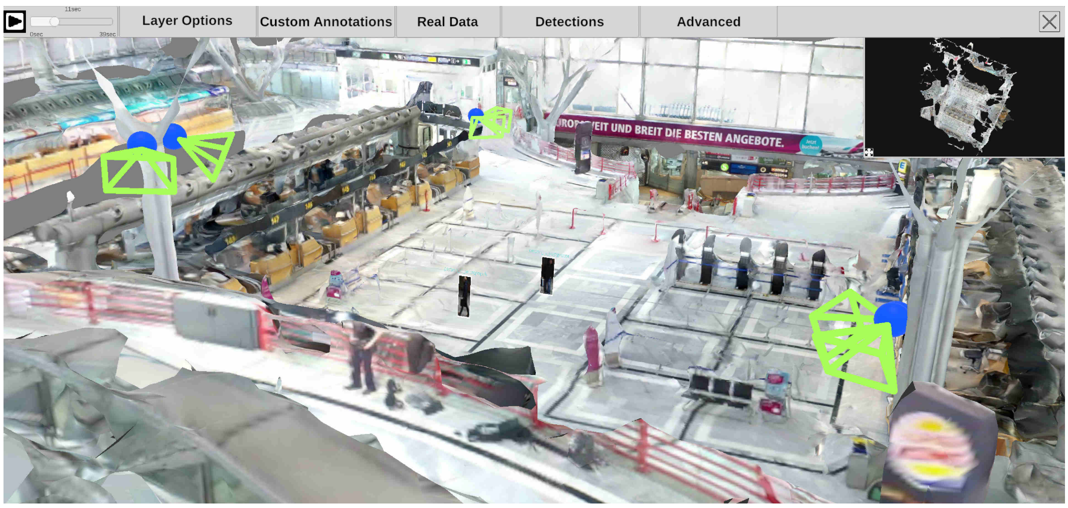

Figure 5.38: Reconstruction of airport in which multiple surveillance cameras are spatially registered. Video streams of the cameras are fed into the system and automatically extracted detections are depicted in the 3D reconstruction in real-time.

Their video streams are processed in real-time in a pipeline for object and person detection. The extracted meta information is continuously displayed within the 3D scene. Security staff can then interactively monitor a single representation of the entire complex, rather than a wall of monitors, showing the footage from a single surveillance camera. Besides, the model would provide a good overview of the distribution of cameras in the complex, making it easier to follow suspicious persons walking past different cameras, even if the cameras' original video recording is monitored. Especially in use cases like this one, however, it must be critically reflected from a data protection and ethical point of view to what extent technical possibilities should be employed. For example, the tracing and re-identification of a person could already constitute a massive encroachment on a person's rights.

5.2.5.4 Mission Planning and Training

Another use case for the presented demonstrator would be applying 3D reconstructions for mission planning and training scenarios. For example, in the case of a hostage situation in a university, the available video and photo footage of its complex could be used to create a reconstruction. Additional material collected during recon missions by robots or drones can be used to improve the 3D model. Special police forces could then use this reconstruction to get a picture of the surroundings and plan strategies. For example, animated annotations can be used to sketch possible ways to enter the building and free the hostages (see Figure 5.39).

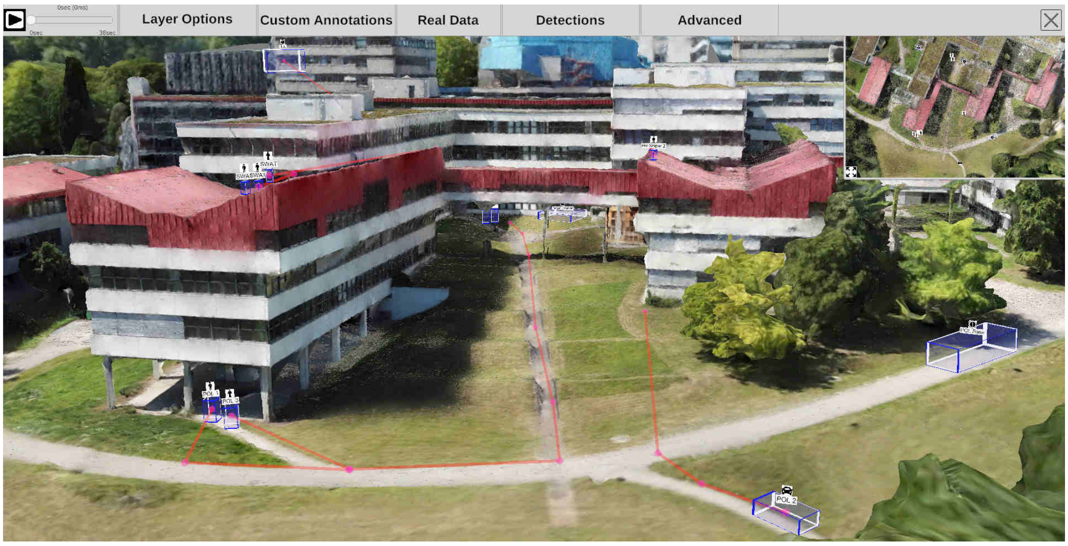

Figure 5.39: Exemplary reconstruction of the environment for strategy planning in police operations. The demonstrator creates a static mesh from drone recordings. The planned movement of police forces can be sketched in it.

Furthermore, police forces could use models of past missions or create new 3D models of training environments to train their personal. Employing VR could be of particular benefit in such training scenarios, as trainees can enter the given surrounding and perceive them more realistically. Previous research points to advantages of virtual reality in terms of memorability [245], spatial navigation [319], orientation [313], learning performance on 3D models [145], and the understanding of complex geometries [429], which could also lead to an overall better training effect in the given context.

5.2.6 Discussion

5.2.6.1 Limitations

The current version of the demonstrator has different limitations. Like other state-of-the-art 3D scene reconstruction algorithms, the current algorithm is sensitive to low-quality input material. For example, differences in illumination in images taken from the same location may mean that no commonalities may be found between the two sources, leading to the inability to identify spatial links. The image quality is also of great importance. Artifacts caused by motion blur in videos or low image quality, e.g., due to poor illumination during night shots, have a significant impact on the reconstruction quality and can make it impossible to register cameras. Therefore, it is not guaranteed that all videos that record a specific location are also registered in space and thus revealed in the visual analysis process. Additionally, cameras may be incorrectly registered due to a confusion of crucial points. Both cases pose a threat to a possible decision-making process during visual analysis. Therefore, it is essential to verify automatically calculated and extracted information and consult the original data for the final decision making. Although the current framework provides links to original data, the convenience provided may discourage additional verification steps.

In the current approach, the reconstruction can be recognized as a non-realistic estimation of the environment. However, with improving reconstruction algorithms, the environment's quality and level of realism might align more and more with real-world experiences. Like this, errors in the reconstruction might be accepted without further verification. Therefore, the currently already provided functionalities to quickly open original video and image footage will be important in the future.

We presented our tool's functionality for navigation through a 4D reconstruction in VR. Although this offers several advantages, it also has drawbacks. Some users are prone to simulator sickness and become nauseous after a short time of immersion, limiting the target group of potential end users. Moreover, even though the illusory reality looks spacious, the physical interaction space is usually limited, resulting in a small area where users can actually walk naturally. Movement-compensating techniques such as virtual teleportation must be used to cover greater distances in VR, but can negatively influence the perceived presence and orientation [313].

Beyond that, the demonstrator presented is intended to convey the underlying concepts and not represent a ready-to-use prototype. Therefore, it cannot currently scale to hundreds of input videos. The preprocessing should be outsourced to a GPU cluster, enabling highly parallel processing to process more massive amounts of input data. This would lead to an almost linear reduction in computation time since most high computing power steps can indeed be parallelized.

5.2.6.2 Ethical Considerations and Legal Aspects

Naudts and Vogiatzoglou state in "The VICTORIA Ethical and Legal Management Toolkit" that every application of new technology should consider several general ethical principles [264]. The proposed approach requires readily available imaging data. Of course, this data needs to be gathered in a lawful and ethical way. For example, this could include locations such as airports, train stations, or other public places where CCTV cameras are already deployed. To create a proper reconstruction, additional imaging material from a moving camera sensor is necessary. It is best to achieve a high-quality reconstruction if no persons or other moving objects obstruct the view. Thus, no personal data is required for the reconstruction.

The following list of general ethical principles [264] need to be discussed that concern the proposed demonstrator. Beneficence is a principle stating that new technology should improve the individual and collective well-being. Our approach is designed to improve the way users can access 4D imaging data from multiple camera sensors in a more intuitive manner. In general, this measurement can improve the overview of complex scenes, such as an airport, enabling security personnel to detect important events such as an imminent threat and eventually respond faster. In the case of a crime scene reconstruction, a 4D scene may enable criminal investigators and legal experts to improve the decision-making, and the trial process as the specific spatial properties of a crime scene can be investigated exploratory and immersively using virtual reality. This is also relevant in the principle of the right to a fair trial. The same argument is also valid for our use case of mission planning and training, potentially saving lives.

On the contrary, the principle of non-maleficence states that new technology may not be exploited to harm human beings. Our most significant concern here is that our approach may lead to an increase in surveillance, as it allows humans to maintain the overview of a scene even if more camera sensors are being added. On the other hand, deploying CCTV cameras must be aligned to the law and is heavily regulated in many countries.

Justice and fairness are heavily discussed topics in scientific communities [272] and politics [299]. This topic is related to non-discriminatory AI. Both topics are notably complex, and no general solution seems to be available. Naudts describes how this is also reflected in regulations as they are complex and multilayered [263]. The modular system employed in the presented demonstrator also includes the detection of objects and persons. The demonstrator merely receives a class, bounding box, and the respective frame plus additional metadata such as uncertainty values. Our tool displays all available data and does not filter, for example, by uncertainty values to mitigate the problem of fairness. However, the problem that certain aspects of a scene may remain undetected persists, but it is more unlikely the more cameras and frames are available containing the object. Another optional module is the re-identification if this is enabled, it allows the user to track objects through the scene; for example, by visualizing the paths and lifelines. Such modules may require the use of biometric data such as detecting persons by their faces. Therefore, the lawful applicability must be ensured. However, as Kindt argues, clearer rules regarding the use of biometric data are required [188]. This is also heavily affected by GDPR. Our approach is robust to deal with, for example, with blurred faces. The data must be prepared before it is loaded into our tool.

The principle of autonomy states that humans must remain in control over important decisions affecting themselves and others. In the research field of visual analytics, this is also known as the human being, the ultimate decision-maker. In a criminal investigation, interactions with a tool relevant for decision-making must be tracked and presented at court in combination with the findings [178]. We envision our approach and tool as an alternative view for 4D scenes. It does not automatically make decisions except for the reconstruction itself. The tool furthermore always allows the operator to access and view the original data and relevant metadata. This measurement also complies with the principle of explicability and eventually increases the operator's trust in the system.

5.2.6.3 Future Work

New techniques and approaches are continually being developed, which are improvements of current steps in our preprocessing pipeline. For example, new, faster, and more accurate ways are developed to detect objects and persons in videos, re-identify them in other frames or videos, and extract metadata from them. The presented demonstrator is based on a modular design that allows the continuous adaptation to technological progress.

In the future, outsourcing the preprocessing pipeline to a GPU cluster should increase processing speed and facilitate the analysis of large amounts of videos. We also plan to improve various steps in the preprocessing procedure, such as the extraction of meta information. Although skeletal data can be easily extracted from detected persons, they are not yet classified for further analysis. Another next step would be to apply behavior classification networks to the skeletal data. As a result, the skeletons would be labeled with tags describing their current state within a particular frame, such as “walking” or “sitting”. We also plan to improve the inter-video re-identification of detections to calculate 3D locations of detections, merge 3D point clouds of the same detection, and create 3D avatars for detected persons.

Currently, the reconstruction of the static 3D environment and the cameras’ spatial localization are computed in a single step. It would be advantageous if sources could be added incrementally to an existing 3D reconstruction. In this way it might even be possible to register a moving camera on-the-fly in the static reconstruction and enable image-based position tracking.

Furthermore, we plan to extend the applicability of the tools for collaborative investigations. Currently, only one person at a time can enter the virtual environment. In the future, however, the remote collaboration of several users in virtual reality should be supported. The option to enter the same virtual environment and explore the 4D reconstruction interactively could improve the dialogue between remotely located participants due to the improved communication basis [64, 257]. Future research should assess the possible benefits of remotely co-located collaboration in this context.

In addition to a quantitative evaluation of immersive analytics for the interactive analysis of 4D scene reconstructions, we also plan to systematically evaluate the overall system by assessing its performance for various analysis tasks and comparing it with alternative approaches. For selected databases, users will try to solve specific tasks and extract high-level information from the data. Tasks could range from basic questions such as “How many different people are visible in all videos?” or “Who drops a suitcase when and where?” to more complex analysis tasks such as “Find the person who threw a bottle, trace back where he/she came from, and extract a frontal image of the person’s face”. Besides performance and task completion times, additional measures such as usability and workload will be taken into account.

5.2.7 Conclusions

This work introduces a framework for the interactive, visual analysis of mass image and video data. The framework consists of a modular preprocessing pipeline that prepares a highly unstructured and heterogeneous bulk of digital footage for later display. Besides the temporal and spatial registration of the sources in a static 3D reconstruction of the corresponding environment, meta-information is extracted for each video. Therefore, the user can spatially and temporally explore the data while maintaining an overview of all materials. The main advantage of this approach is that all information is presented in a shared visual context, which reduces users’ mental effort to link different sources. Besides, the framework enables immersive exploration of the data space in VR, allowing the analyst to “enter” the 4D reconstruction and search it more naturally. To illustrate the versatile applicability of the framework, four use cases for different application areas such as crime scene investigation, real-time surveillance, mission planing, and training scenarios were presented. Initial qualitative assessments by criminal investigators underline the potential of using virtual reality for the exploration of 4D reconstructions, as it fosters spatial understanding, allows more intuitive ways of collaboration, and enables remote inspection of crime scenes in a natural way.

6 Conclusions and Perspectives

Contents

After presenting several examples of how to assess the applicability of VR for data visualization using all three strategies, this concluding chapter reflects on the research area at a meta-level and places the body of research presented in the broader context of its research objectives. The reflection is based on our opinion paper [201]. In addition, key contributions are summarized, future research directions are outlined, and perspectives on immersive analytics are discussed.

6.1 Reflection: The Value of Immersive Analytics

In recent years, research on immersive environments has experienced a new wave of interest, and immersive analytics has been established as a new research field. Every year, a vast amount of different techniques, applications, and user studies are published that focus on employing immersive environments for analysis procedures. Nevertheless, immersive analytics is still a relatively unexplored field that needs more basic research in many aspects and is still viewed with skepticism. Rightly so, because in our opinion, many researchers do not fully exploit the possibilities offered by immersive environments and, on the contrary, sometimes even overestimate the power of immersive analytics. Although a growing body of papers has demonstrated individual advantages of immersive analytics for specific tasks and problems, the general benefit of using immersive environments for effective analytic tasks remains controversial. In this section, we reflect on *when* and *how* immersion may be appropriate for analysis and present four guiding scenarios. We report on our experiences, discuss the landscape of assessment strategies, and point out the directions where we believe immersive analytics has the greatest potential.

Immersive Analytics is the research on analyses concerned with the “use of engaging, embodied analysis tools to support data understanding and decision making” [56]. Such ‘engaging tools’ include augmented reality and virtual reality devices. Over the last decade, IA has gained attention in the scientific community, particularly in the areas of visualization and human-computer interaction. There have been repeatedly times in the past when research on immersive environments has been particularly intense [120]. The recent surge may be due to the technological advancements in consumer-ready head-mounted augmented reality and virtual reality displays, as well as the stronger inclusion of the analytical process in such environments. Even though more and more research is being produced each year, the field as a whole is still relatively unexplored. Fast-paced technological progress means that research is targeted at research subjects that are rapidly changing. Findings and conclusions that apply to one device may not be applicable to another device that has a higher resolution, a wider field of view, or any other change that improves the immersive experience.

Although immersive analytics (IA) aims at multisensory interfaces, the focus is often on vision. Most researchers now agree that IA is not a panacea that overcomes all issues associated with 3D visualizations on screens and makes unfavorable 3D visualizations suddenly useful. The underlying drawbacks of these 3D visualizations [258], such as occlusion, remain even when viewed in an immersive environment. Nevertheless, we are under the impression that many IA studies are conducted with abstract 3D visualizations, such as scatterplots, without comprehensive justification. In some cases, even 3D visualization variants that are generally believed to perform worse than 2D counterparts are compared based on their performance on different media such as screen vs. AR/VR. That is, many studies use abstract 3D visualizations in immersive environments which have already been shown to perform poorly in the past, rather than shifting the focus to other visualizations and application domains that are much more likely to actually lead to advantages in immersive environments compared to classic 2D screen setups.

This and similar circumstances have led us to question whether many current efforts are heading in directions that do not exploit the full potential of immersive environments. While it is legitimate to revisit and reevaluate previous findings with new devices, the focus should be on approaches that promise the greatest potential in the extended design space. AR and VR offer much more than just a medium for viewing 3D visualizations, for example, by greatly expanding the design space in terms of multisensory interfaces, interaction, navigation, and collaborative aspects. Similar to Dwyer et al. [102], we define the term immersive analytics very broadly and regard it as an interplay of analytics, visualization, interaction, and multisensory experiences.

Given the previous hype periods for certain technologies such as VR and AR, we think it is important to mention the fundamental differences of the current surge, such as the wide availability of affordable, high-quality devices, and the existence of whole software ecosystems and communities which greatly simplify the implementation effort. However, we also like to point out potentially remaining obstacles. These include limitations of both the medium (hardware/software) and the human user. Expanding the range of data representation characteristics, e.g., to a multisensory 3D representation, is more prone to emphasize group differences and perceptual deficiencies of the human user than the limited classical 2D visualizations. For instance, stereo blindness or movement deficits may affect analysis or data interpretation. Wearing tethered VR goggles for several workdays could have strong effects on human health and well-being, and thus be prohibited for use in certain work environments.

Based on these considerations, our driving question is the following: *Why, when, and how does it make sense to use AR/VR for analysis tasks?* First and foremost, we want to make the reader aware of (1) the fact that IA does indeed extend the design space of classic visual analytics, (2) the plethora of opportunities for developing new analysis, visualization, and interaction techniques, (3) potential risks and common pitfalls, and (4) underexplored, yet promising research areas.

Skarbez et al. [343] recently outlined a general research agenda for immersive analytics. We complement their line of argumentation by presenting four guiding scenarios that illustrate where we believe some of the greatest potential for immersive analytics applications lies and discuss the value of IA. These scenarios were derived from the experience and discussions among the authors. We conclude this section with a summary of lessons learned, including references to promising research gaps, appeals to avoid common pitfalls, and general remarks on the topic.

Four Guiding Scenarios

We take a look at four scenarios where we believe IA has the greatest potential. The list is not exhaustive, and there are certainly additional directions that are generally promising.

I - Situated Analysis

Scenario Augmented reality fosters the presentation of situated visualizations, that is, the embedding of visualizations in the real environment close to the object of their content. Due to the proximity of the information to the object it refers to, the connection can be easily understood. Embedding visual information directly into its physical context is usually not possible with classical user interface setups. The approach implicitly follows the principle of ‘details on demand’, as the data space is continuously filtered for information that is displayable in the user’s field of view. Thus, only information that is potentially interesting to the user at a given location is displayed. While glyphs could serve as initial visualizations to provide a good overview, users could be allowed to interact with them to dive into even more details.

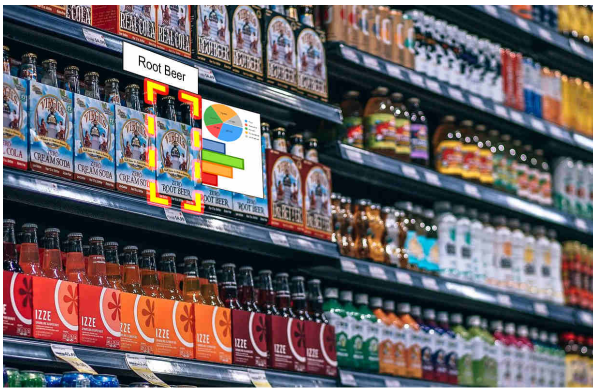

Figure 6.1: Situated visualizations that display custom quality scores for each product on a grocery store shelf.

Examples A common example is the display of nutritional information as a bar chart or glyph visualization above each item in a grocery store, as illustrated by ElSayed et al. [107]. An example of what this might look like is shown in Figure 6.1. Also quite popular is the dynamic placement of labels in AR space. For example, Zollmann et al. [433] use AR to place labels next to buildings to provide users with additional information about their surroundings. Additionally, situated visualizations could also be used to support user navigation, for example, by displaying a trail on the floor that leads to the desired shelf in a library.

Reflections Situated visualizations and therewith facilitated situated analyses are certainly a big selling point of augmented reality. The possibilities are almost endless once the technology is fully there and AR is widely accessible and used by the public. This ranges from advertising to informative and supportive visualizations to situated visualizations for analyzing real-world environments or objects. Of course, this also comes with challenges. For instance, an increasing degree of augmentation can lead to neglecting the real environment and to sensory overload. Additionally, users must trust the program, as it can direct the users' attention and influence their perception.

II - Spatial Data and Spatial Tasks

Scenario Analytical procedures that deal with the examination and analysis of spatial data can benefit from immersive environments. Spatial data often has an inherent spatial mapping in 3D space, while for a representation in 2D, some sort of transformation or abstraction must be applied. Of course, whether retaining the original structure is a significant advantage over abstraction depends on the individual data and tasks. However, especially for spatial tasks or when a spatial context such as the natural environment has to be integrated into the motion analysis, the deployment of 3D visualizations can be useful. Once there is a clear motivation for using 3D, additional benefits of 3D visualizations can be exploited in immersive spaces. For example, anyone who has worked with 3D modeling software knows how difficult it can be to navigate in 3D space or to select a specific 3D region using keyboard and mouse. In immersive spaces, such tasks could be achieved more easily by providing direct interaction capabilities in the 3D space since no translation from the 2D input space to the 3D space is required. In addition to the benefit of an expanded, multimodal interaction design space, previous work has shown potential advantages of stereoscopic and immersive devices that could also be exploited, such as enhanced learning performance [145], memorization [300], spatial understanding [12], and orientation [334].

Examples Hurter et al.'s FiberClay [166] is a framework for exploring 3D aircraft trajectories in a VR environment (see Figure 6.2). Exploring the trajectories in an immersive environment allows the analyst to make use of intuitive spatial interactions, e.g., for selection, while preserving the original shape of the trajectories. Additionally, stereoscopic vision helps to distinguish different trajectories and estimate their depth. There are further examples from other domains, for instance, a scenario from the medical domain for analyzing brain scans where a brain or associated data is interactively investigated in 3D space [99, 173, 270]. Similarly, in biology and chemistry, the same approach can be used to visualize and analyze microorganisms and molecules [269, 271].

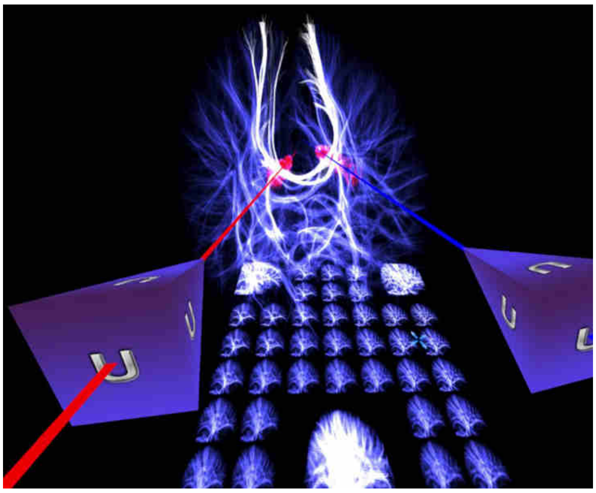

Figure 6.2: Analysis of 3D trajectories in virtual reality [166]. Image courtesy of Christophe Hurter.

Reflections Particularly for the pairing of spatial data and spatial tasks, the use of AR/VR is often promising. In those cases, there is a clear motivation for visualizing in 3D, and immersive spaces offer advanced interactions to facilitate spatial tasks in 3D. However, the need to visualize spatial data in 3D should be confirmed and first compared with 2D alternatives [29]. For this scenario, we see the biggest challenge in resisting the temptation to rely on immersive, spatial solutions when better 2D alternatives exist. Additional challenges with the analysis process itself include difficulties in designing interactions in 3D space or hardware limitations such as a too low resolution to read text labels properly.

III - Collaboration

Scenario Immersive environments offer various advantages when it comes to collaboration. In our opinion, the biggest opportunity for improvement lies in remote collaboration. By using AR/VR, multiple users who are in different physical locations can meet in a shared virtual environment. This gives them a common visual grounding to support their discussion while allowing them to use familiar communication aids such as gestures, facial expressions, and simplified verbal expressions related to relative positions in space (e.g., "here", "left"). Of course, the extent to which this corresponds to real-world, co-located collaboration experiences is highly dependent on the technical implementation, such as the photorealism of avatars and the achieved embodiment in one's own avatar, e.g., through the perception of one's own body, the provision of a large interaction design space, and haptic feedback.
Another advantage of remote collaboration is scale. For instance, while usually only a limited number of people can stand around a ship's engine, in the virtual environment a large number of people can simultaneously observe the 3D model of the engine - even from the exact same location if their avatars are rendered invisible. In addition to the advantages in terms of practicability, interaction, and communication, other social aspects could be exploited in the future. Since user avatars can be designed arbitrarily, it is possible to overcome social inequalities by designing them neutrally in scenarios where this is an issue. There are also possibilities for co-located collaboration, some of which overlap with those for remote collaboration scenarios. For example, when viewing a 3D visualization, all users can simultaneously investigate the same visualization while constantly seeing where others are. This can improve communication compared to a setup where all users are observing the same visualization but on separate screens.

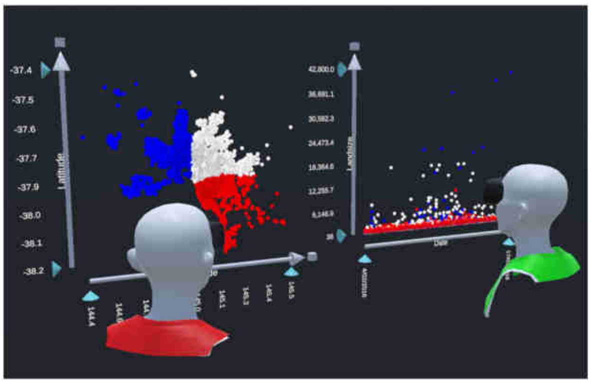

Figure 6.3: Collaborative VR environment for the analysis of abstract 2D and 3D visualizations [221]. Image courtesy of Benjamin Lee.

Examples Lee et al. [221] presented Fiesta, a system for collaboration in physically co-located VR environments. Multiple users can join a shared VR environment to analyze abstract data visualizations together. The visualizations presented are not necessarily in 3D, and the VR environment can be used simply as a platform for collaboration without changing the familiar visualization basis. Another use case could be the deployment of immersive environments in teaching scenarios. For example, a real classroom could be replicated in a virtual environment so that students in remote locations can participate in digital lessons and experience them similarly to real classes. Moreover, the use of AR/VR can improve social interactions and communication. In our opinion, the Corona pandemic in particular has shown that video chats cannot compete with face-to-face meetings in many respects. Realistic imitations of real meetings using immersive environments could therefore have a lot of potential.

Reflections In our opinion, collaborative analysis tasks, in particular in remote collaborations, can certainly benefit from IA. Currently, most examples are avatar-based VR applications. There are few examples of AR being deployed for this task, and there are several issues that need to be addressed for better AR-based remote collaboration. For example, AR applications share only a fraction of the overall environment since all collaborators have different real environments, and the display of avatars is a barrier because many AR devices are gesture-based and therefore do not have steady information about the position of the arms, making it difficult to display avatars correctly. One of the biggest challenges related to immersive collaboration is its susceptibility to impaired communication due to unwanted artifacts. For example, imperfections in copying participants' gestures and facial expressions can lead to major misunderstandings among collaborators. Sometimes the deliberate suppression of nonverbal expressions can be beneficial. In addition, the technology is not yet widespread to be used in everyday life and only participants who have the right hardware can collaborate.

IV - Presentation

Scenario Immersive environments can be appropriate for simply presenting information - but in a more engaging way. The use of the relatively new and unfamiliar environment is associated with higher levels of excitement, engagement, and entertainment [177, 256, 297]. Such effects certainly help users to keep their attention and internalize information. However, it is not yet clear whether the effect will diminish as the technology becomes more familiar and the 'WOW' effect wears off. Another potential benefit of using AR/VR is that it can help users relate familiar measures like distance, speed, or height to themselves, leading to better estimates of their absolute values. For example, when perceiving the 3D model of a house, it is easier to estimate its actual size without reference scales in VR than on the screen [220]. The goal of presentation is to convey information as completely and sustainably as possible. Previous studies have shown that immersive environments can support users' (spatial) memory due to more engaging illustrations and spatial anchors (e.g., [199]). Therefore, if this feature can be exploited in a particular presentation scenario, this could be a good motivation for using immersive environments.

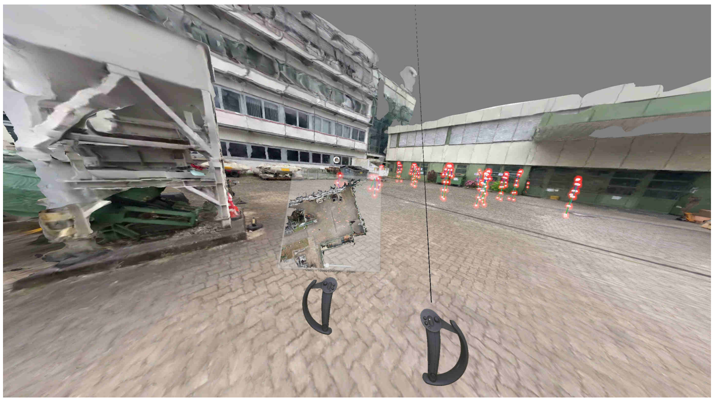

Figure 6.4: A reconstructed crime scene is used to vividly present the sequence of events to a court jury [204].

Examples For abstract data, spatialization can be useful to exploit the properties mentioned above. For instance, Zenner et al. [428] presented a way to represent circuit diagrams as 3D landscapes that can be explored in a VR environment. The authors concluded that although vivid presentation increased user interest, it had no impact on model understanding performance. Another example is the remote access to reconstructed environments, such as museums, construction sites, or excavation areas. Users can walk through the virtual reconstruction of a real environment without having to physically move. For example, in Figure 6.4, a reconstructed crime scene is shown in a VR environment to vividly convey the course of events in a court trial [204].

Reflections While in conventional screen-based analysis environments a lot of effort is put into increasing the level of engagement through clever user interface design or even gamification, this already seems to be a side effect in immersive analytics. However, it may well be that the effect diminishes with increasing usage. In addition to potential benefits in terms of higher engagement, improved absolute value estimation, and memorization, immersive environments could also be suitable for information presentation during remote site inspections, lectures, or corporate presentations. The three biggest challenges in this scenario are availability, usability, and accessibility. Availability refers to the fact that AR/VR is not yet 'common enough' and only a small portion of the population owns AR/VR hardware. By usability we mean the circumstance that AR/VR is still unfamiliar to many people and many different interaction designs exist that are often difficult to grasp. Finally, by accessibility, we refer to the challenges inherent to new interaction designs and sensory stimulations which are not accessible to some people.

Lessons Learned

As illustrated, there are several scenarios where we see great potential for immersive analytics applications. In the following, we outline lessons learned, address best practices, and discuss common pitfalls.

Immerse when it adds value.

Repeatedly, we have come across examples where immersive environments were used seemingly for no reason - just because the technology was new and available. However, when using immersive hardware, there should at least be a hypothesis that promises added value. The actual use of the technology should then be guided by the objective assessment of the added value. The extent to which AR/VR capabilities are exploited must also be carefully considered. It may not make sense to force the user to walk for spatial navigation or even to perform all analysis steps in an immersive environment just because it is possible. For example, if the IA approach is only beneficial for a specific subtask in an analysis procedure, it may make sense to use hybrid environments where only part of the analysis is done in AR/VR and the rest on a traditional screen.

IA is not the Holy Grail of 3D visualization.

A particularly controversial issue is the visualization of abstract data in immersive environments. Although reading 3D visualizations is improved in some respects when perceived in immersive environments, most of the drawbacks of 3D visualizations remain. For instance, occlusion, shifted baselines, depth distortions, and the difficulty of estimating and comparing certain visual variables, such as volume, remain major problems. Thus, even if the particular evaluation can show that AR/VR improves the analysis with the 3D visualization compared to a screen-based setup, it says nothing about the overall merit of AR/VR, as more powerful 2D alternatives for the screen were simply left out of the comparison (see also the “Straw Man Comparison” pitfall outlined by Munzner [258]).

However, there are certainly specific application areas where it may be useful to spatialize abstract data in order to take advantage of the aforementioned benefits of immersive environments, such as improved spatial understanding, orientation, memorization, or depth perception. For example, in the comparative analysis of 3D distributions of abstract data, the immersive, spatialized 3D variant with superpositioned 3D heatmaps was superior to the juxtapositioned 2D variant in certain tasks [198]. As shown in Figure 6.5, the vertical layout combined with the encoding of values on heightmaps facilitates the comparison of the two distributions. The user can slide one distribution through the other to identify correlations, offsets, and general trends. Another example where immersive environments can be useful for abstract data is the integration and exploration of abstract and spatial data, which has been discussed for some time, for example, for applications in life sciences [185].

In this sense, the use of AR/VR should not be the only motivation for 3D visualizations. It may be that IA makes 3D more feasible, but the associated disadvantages must be outweighed by advantages to justify the deployment of the 3D visualization in question. At this point, it is worth mentioning that IA goes beyond 3D visualization, and its added value can also be drawn for 2D visualizations from other aspects, such as multisensory feedback, enhanced interaction modalities, collaboration opportunities and so on.

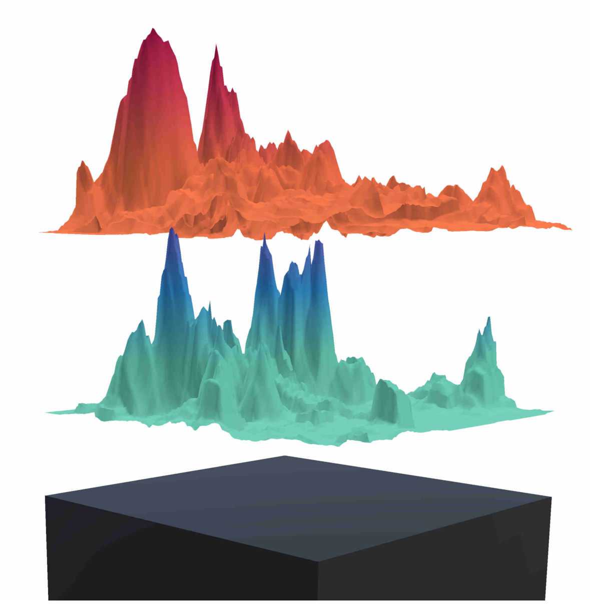

Figure 6.5: Abstract data is spatialized, displayed as stacked 3D heightmaps, and observed in virtual reality for comparative analysis [198].

Assess the value of deployed AR/VR environments.

Assessments of added value, as they are often used in practice, can be divided into three main groups. The first and weakest assessment is that by logical reasoning. In this case, certain properties of and circumstances in immersive environments are argued to be beneficial for a certain analytical procedure based on logical arguments and evidence from related research. Even though this approach does not generate evidence itself, it provides connecting points for new hypotheses that can be asserted in user studies. The second form of assessment is property evaluation. A specific aspect is singled out and compared across different media; usually immersive and non-immersive analysis scenarios are compared. An example would be a study comparing the memorability of users when observing a visualization on a screen and in VR. While this may provide the most reliable and substantiated evidence, it could well be dependent on many factors that would not apply in a particular application scenario. The final group of assessment is concerned with the (often) qualitative evaluation of a technique of application as a whole. This form of assessment can identify the overall performance, as well as some advantages and disadvantages of a particular IA system over state-of-the-art analysis procedures. But as often only one system is evaluated without direct comparison, and because of many independent variables, the exact reasons for differences in performance are often difficult to determine.

We argue that all three types of evaluation have their right to exist. While the first approach provides initial ideas and new hypotheses, the second approach can quantitatively explore potential merits at a very detailed level. Their applicability and usability for a particular analysis use case can then be assessed by means of the third form of evaluation. For the second type of assessment, in particular, it is important to ensure that a fair comparison takes place. For example, in most cases, it does not make sense to assume the use of 3D visualizations when much more powerful 2D visualizations exist for the given task and then compare the performance of users working with them on screen and in VR. In case it is assumed that the use of an immersive environment overcomes the disadvantages of the 3D visualization, the 3D visualization in VR could be compared with the best possible 2D visualization on the screen.

Keep Going.

Immersive analytics is still a relatively new field that lacks a broad scientific fundament. For instance, often criticized but not sufficiently addressed is the lack of established analysis environments, authoring toolkits, and standards for immersive analytics. There have been advances in the compatibility of development frameworks such as Unity or UnrealEngine. This has reduced the effort required to create new applications for immersive devices compared to previous VR eras, such as during the 90s, where such development endeavors needed to be much closer to the hardware. However, there is still no end-user-ready visualization system like Tableau for direct manipulation analysis of data, neither is there an established library like D3.js for a unified way to create custom visualizations in immersive environments. Likewise, with regard to interaction modalities, no golden standard – similar to the iconic duo of mouse and keyboard for PCs – has yet emerged among the many options. Every single AR/VR device manufacturer relies on individual controllers and input modalities. Additionally, rapidly evolving hardware leads to the need for continuous re-evaluation. Findings that apply to a CAVE VR environment from the 80s do not necessarily apply to modern HMD VR setups. As there are many different areas of potentially very useful applications for immersive hardware, even away from immersive analytics, we expect the technology to grow in popularity, familiarity, and availability over the long term. And this could make it even more attractive for everyday IA procedures.

Bottom Line

We believe that there is much potential for immersive analytics and that there are ample opportunities for research in this area. In this section, we presented four guiding scenarios to which we attribute great potential of immersive analytics: situated visualizations, spatial data analysis with spatial tasks, collaboration, and presentation. In addition to examples and justifications for our proposals, we also reflected on the overall situation in the field and pointed out common misconceptions and – in our opinion – best practices. While in this article we focused on immersive visualization, the field of immersive analytics is much broader and has much potential in other directions as well. For example, the involvement of different senses such as through sound or haptics, opens up a whole landscape of different design opportunities for analytic processes. For each individual aspect, the new possibilities bring new challenges, and it is up to research to determine the added value of immersive analytics for a given combination of data, task, and user. In the past, there have been several research hypes of immersive technologies that promised great changes – which never occurred to the anticipated extent. Even if there are technological opportunities and improvements, the technology must first be accepted by potential user communities. While it is still unclear when and how immersive technologies will become standard tools for data analysis, there are already strong indicators such as studies with convincing evaluations that show the potential as well as the availability of development platforms, but also software and hardware sales, that let us expect that these technologies might be here to stay this time.

6.2 Summary of Contributions

The guiding research question of this thesis is "How can the applicability of virtual reality for data visualization be assessed?". This chosen scope of 'assessment' is not strictly limited to evaluation strategies and approaches but generally aims at ways to judge if and how the deployment of VR is reasonable for a particular application. To answer the research question, the assessment spectrum was divided into three types of strategies, which were discussed in three chapters, each illustrated by at least two examples. While the discussion of the entire set of assessments and their categorization leads to findings on a meta-level, the greatest contribution of this thesis lies in the individual assessments, which present diverse findings in different research directions.

Main findings of assessments.
In the following, the main findings of the individual assessments are summarized and grouped by their subject of conclusion (*Why / How*). Additionally, the assessment strategy used to arrive at the respective conclusion is pointed out.

Why deploy VR?
We presented an assessment that argumentatively justifies that in specific use cases, it can be useful to deploy VR for analysis validation, as it transports the user into the analyzed scenario while trying to convey the scenery as realistic as possible. This makes discrepancies and inconsistencies with real-world experiences quickly apparent to the user. Similarly, another assessment presents a collection of arguments why VR is, at least in theory, not inferior to conventional media for performing visual analysis procedures, while offering features that are not available in conventional screen-based setups, such as remote co-located collaboration, high portability of hardware, and large visualization design spaces. Both assessments can be classified under the first assessment strategy, as they attempt to justify the deployment of VR with arguments that are not fully based on grounded theory. In two controlled user studies, we applied the second strategy and evaluated two different human factors in two different ways. First, we evaluated the impact of immersion on a conventional cluster identification task in scatterplots in which we manipulated the environment to be more or less immersive. Results showed, among other things, that immersion has a positive effect on memorization and orientation. Second, we evaluated the impact of different visual navigation aids on orientation and manipulated the visual tool provided. We found that the optimal choice of navigational support is highly dependent on the task. For example, a minimap is best for keeping track and returning to known targets. Last but not least, we presented two examples of the third strategy and evaluated two applications. The first evaluation showed that VR leads to improved performance on various comparative analysis tasks on 3D heatmaps, while it performs worse on other tasks. Thus, while we were able to demonstrate advantageous aspects of VR, its application for the comparative analysis of heatmaps needs to be carefully weighed depending on the specific task. The second demonstrated application was a prototype for the interactive and immersive analysis of a crime scene reconstruction. Qualitative expert feedback revealed interest in the novel approach and attributed a high value to the possibilities it opened up.

How to deploy VR?
Several findings, mainly in assessments of the first and second strategies, provide guidance on *how* virtual reality should be implemented. In the first chapter, we discussed the level of realisticness that

is desired in VR environments. By weaving a framework of arguments, we came to the conclusion that it is not always reasonable to strive for ultimate realisticness by replicating every detail of real environments. Some visual metaphors should be avoided and replaced with less realistic, more efficient alternatives. Quantitatively evaluating the impact of immersion on cluster identification tasks, we compared different visualization spaces and concluded that a spatially restricted tabletop visualization is superior in many ways to a room-sized visualization that encloses the user. The tabletop visualization increased overview capabilities, promoted the maintenance of orientation, and minimized self-occlusion. Results of the second quantitative user study on orientation provided insights into how navigational tools for VREs should be designed. Depending on the task, optimal visual support can be chosen. For example, for tasks that require high retracing skills of users, a minimap might be more favorable than a heatmap on the floor, which in turn is better suited for tasks that strive to optimize spatial exploration efficiency by minimizing redundant visits.

Main contributions and future research directions.
Finally, the main contributions besides the findings from the literature survey and the individual assessments are briefly summarized along the structure of the thesis and the respective future perspectives are addressed.

Establishing the argumentative basis for the use of VR for data visualization.
After an overview of related work in chapter 2, chapter 3 first addressed the argumentative basis for conducting research in IA. The first type of assessment strategy presented therein builds on the collection of arguments from the literature to argue for and against the use of immersive environments for analysis purposes and to show how to apply them. The first section (section 3.1) illustrated an example of such an assessment by arguing that VR could be beneficial for certain analysis verification steps. Subsequently, the second section (section 3.2) continued the discussion at a higher level. The central idea was: What would happen if we replaced the conventional medium 'screen' in a typical visual analysis pipeline with a VR medium? We presented a step-by-step discussion in which we first assumed an optimal VR environment, which was then gradually approximated to realistic conditions. From this we concluded that – theoretically – VR is in no way inferior to conventional media and even offers unique opportunities. However, due to technological shortcomings and typically divergent uses of immersive environments, this is often not the case, and it must be assessed for each individual case if the deployment of VR is appropriate. In the last section (section 3.3), the focus was shifted from the question of 'if' it makes sense to deploy VR for data visualization to the question of 'how' the VR environment should be implemented. In this section, we discussed whether or not it makes sense to replicate the real world as much as possible in the VRE. We found that it is not generalizable how much replication of the real world is desirable, and that whether and to what extent such replication can be beneficial for a particular application depends heavily on the scenario. While our findings are reliable and universal, and motivate further research in this area, they also highlight some mildly supported hypotheses that require additional investigation to solidify the key messages. Future work could extend the presented framework of arguments, discuss further guidelines, but also, more importantly, follow up on the generated hypotheses and perform quantitative evaluations.

Probing the properties and characteristics of immersive environments.
In the next chapter (chapter 4), we focused on the second assessment strategy. In this approach, controlled studies of properties, characteristics, and conditions of immersive environments are conducted to gain insight into the rudimentary interplay of cause and effect. For example, the goal could be to evaluate the influence of immersion, a unique property of immersive environments, on memorization. If a strong effect is found, these properties could be systematically used in future applications to take advantage of their impact on users' increased memorization capabilities. We presented two examples of such evaluations. The first one (section 4.1) essentially compared a conventional monitor screen to a VRE for the analysis of scatterplots, using four conditions with increasing levels of immersion. In the quantitative user study, participants completed cluster identification tasks on 2D scatterplot matrices and 3D scatterplots. Our results indicated that as immersion levels increased, orientation and memorization improved and learning curves flattened, while interaction effort and task completion times increased. When designing future analysis environments for similar visualizations and tasks, our results can be taken into account. For instance, if a task on a 3D scatterplot visualization requires a user to have good orientation skills while maintaining the overview of the entire dataset, it might be useful to deploy a VRE with a visualization that is viewed from the outside. The second example (section 4.2) dealt with the evaluation of different visual aids to improve orientation in a VRE. Thus, in contrast to the first example, conditions rather than unique properties of VREs were compared. In a quantitative user study, participants were provided with a specific visual aid to solve several pathfinding and navigation tasks. Our results supported the benefits of using orientation-supporting tools, provided guidelines on when to use which visual aid, and highlighted advantages and disadvantages of each technique.

Probably the biggest limitation of quantitative evaluations like the ones presented is that specific properties are evaluated in a specific context (i.e., environment, technology, interaction modalities). In this context, it is nearly impossible to determine each influencing factor and draw conclusions about its impact. Therefore, the results may not be the same when the same principle is used in a different context. For example, we showed that higher levels of immersion increase participants' recall of 3D scatterplots perceived in a VR environment. Transferring the assumption of increased memorization to other types of visualizations is not straightforward because too many factors change. While the general ideas of our conclusions, such as 'immersion can improve orientation and memorization', are likely to hold for future devices and immersive environments, the details are susceptible to technological advance. Future work can use the outcomes of the present studies as a starting point for new hypotheses to reevaluate the conclusions with new technologies. Furthermore, basic research on the properties and characteristics of immersive environments can gradually produce guidelines and standards that support the development of efficient and effective analysis environments. Particularly with regard to new capabilities in IA, such as multisensory feedback or multimodal user input, many research opportunities remain open and should become the new focus of IA research.

Evaluating applications.
The last main chapter (chapter 5) covered the third type of assessment strategies. That is the evaluation of IA techniques and applications. While the second approach focused on the impact of specific IA properties, this type of evaluation verifies the general applicability of a technique or application to a specific set of data and tasks. In particular, this approach can confirm that certain properties that could

previously only be assessed in a laboratory context far from the real world also apply to real-world applications. For example, if a previous study showed that a particular orientation-supporting technique can improve orientation in an environment that promotes disorientation, such as a maze, the technique could be targeted for use in a more realistic analysis environment. For instance, the technique could be deployed in a visual crime scene exploration scenario to take advantage of induced enhanced orientation. Subsequently, this application can be evaluated and assessed in terms of the desired and expected effects on orientation. Two examples of such evaluations were presented in this chapter. The first application represents an approach for comparing multiple 3D distributions (section 5.1). We hypothesized that enhanced depth perception and direct interaction capabilities would improve user performance on various low-level tasks. Results confirmed a benefit of VR over its non-immersive counterpart for several tasks. A particular contribution of the presented approach is the universally transferable study design with a cross-comparison of medium and visualization. Often, evaluations comparing screens and VREs compare only two conditions: a 2D visualization on a screen with a 3D visualization in VR. Our proposed study design detaches the dimensionality of the visualization from the medium and considers the two as independent variables assessed in a 2x2 cross-comparison. Although this was not expected, our results showed that it is important to examine the impact of each dimension separately, as participants partially performed better in the 2D condition in VR than in the 2D condition on screen. In the second example, an immersive analytics approach for analyzing 4D crime scene reconstructions was presented (section 5.2). It illustrates the versatile application areas of immersive analytics by leveraging intuitive interaction designs, demonstrating opportunities for collaboration, and suggesting hybrid analysis platforms that combine the advantages of screen-based and immersive analysis environments. Future research should continue to develop new techniques and applications for immersive analytics that take advantage of selective properties, which have been positively noted in previous evaluations. Similar to the merits of our presented technique, it can be assumed that there are many more niche applications and tasks that could benefit from immersive analytics and contribute to particularly strong hybrid applications.

6.3 Concluding Remarks

The research conducted was intended to get to the bottom of the question of what the value of immersive analytics is and how it can be assessed. Certainly, this question cannot be fully answered in one thesis, but a contribution could be made. I divided the assessment spectrum into three strategies to determine the value of an IA. For each, I presented several realizations of corresponding assessments. Hopefully, the findings presented can help others (a) evaluate new IA scenarios, (b) decide whether to use VR, and (c) consider new hypotheses for future VR setups. Further, I hope that my research can be used in part to establish sound guidelines in the field of immersive analytics and the design of analytic VR environments.

List of Figures

List of Tables

Acronyms

AR Augmented Reality.

AV Augmented Virtuality.

CAVE Cave Automatic Virtual Environment.

CHI ACM CHI Conference on Human Factors in Computing Systems is the premier international conference on Human-Computer Interaction.

HCI Human-Computer Interaction.

HMD Head-Mounted Display.

IA Immersive Analytics.

ICDP International Conference on Imaging for Crime Detection and Prevention.

InfoVis Information Visualization.

ISMAR IEEE International Symposium on Mixed and Augmented Reality (ISMAR) is the premier conference for Augmented Reality (AR) and Mixed Reality (MR).

IVAPP International Conference on Information Visualization Theory and Applications. IVAPP is part of VISIGRAPP, the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications.

ML Machine Learning.

MR Mixed Reality.

PCP Parallel Coordinate Plot.

SciVis Scientific Visualization.

VA Visual Analytics.

VR Virtual Reality.

VRE Virtual Reality Enviornment.

References

[1] Alabi, O. S., Wu, X., Harter, J. M., Phadke, M., Pinto, L., Petersen, H., Bass, S., Keifer, M., Zhong, S., Healey, C., and Taylor II, R. M. (2012). Comparative visualization of ensembles using ensemble surface slicing. In *Visualization and Data Analysis 2012*, volume 8294, page 82940U, Burlingame, California, United States. International Society for Optics and Photonics, SPIE, doi: 10.1117/12.908288.

[2] Alexander, E. and Gleicher, M. (2016). Task-Driven Comparison of Topic Models. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):320–329, doi: 10.1109/TVCG.2015.2467618.

[3] Allahyar, M. and Hunt, E. (2003). The Assessment of Spatial Orientation Using Virtual Reality Techniques. *International Journal of Testing*, 3(3):263–275, doi: 10.1207/s15327574ijt0303_5.

[4] Andrienko, N. and Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer Science & Business Media, doi: 10.1007/3-540-31190-4.

[5] Andrienko, N., Andrienko, G., and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6):503–541.

[6] Arns, L., Cook, D., Cruz-Neira, C., Arms, L., Cook, D., and Cruz-Neira, C. (1999). Benefits of statistical visualization in an immersive environment. In *Proceedings - Virtual Reality Annual International Symposium*, pages 88–95. IEEE, doi: 10.1109/vr.1999.756938.

[7] Australian Bureau of Statistics (2009). United Nations Statistical Commission and Economic Commission for Europe. `https://webgate.ec.europa.eu/fpfis/mwikis/essvalidserv/images/3/37/UN_editing_glossary.pdf`. From Duplicate 1 and From Duplicate 2 (glossary of terms on statistical data editing - UNDP). Accessed: 14-July-2019.

[8] Azuma, R. and Furmanski, C. (2003). Evaluating label placement for augmented reality view management. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 66–, Washington, DC, USA. IEEE Computer Society.

[9] Azzag, H., Picarougne, F., Guinot, C., and Venturini, G. (2006). Vrminer: A tool for multimedia database mining with virtual reality. In *Processing and Managing Complex Data for Decision Support*, pages 318–339. IGI Global.

[10] Bach, B., Dragicevic, P., Archambault, D., Hurter, C., and Carpendale, S. (2014). A Review of Temporal Data Visualizations Based on Space-Time Cube Operations. In *Eurographics Conference on Visualization (EuroVis)*, volume 00, pages 23–41.

[11] Bach, B., Dragicevic, P., Archambault, D., Hurter, C., and Carpendale, S. (2017a). A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes. *Computer Graphics Forum*, 36(6):36–61, doi: 10.1111/cgf.12804.

[12] Bach, B., Sicat, R., Beyer, J., Cordeil, M., and Pfister, H. (2018). The Hologram in My Hand: How Effective is Interactive Exploration of 3D Visualizations in Immersive Tangible Augmented Reality? *IEEE Transactions on Visualization and Computer Graphics*, 24(1):457–467, doi: 10.1109/TVCG.2017.2745941.

[13] Bach, B., Sicat, R., Pfister, H., and Quigley, A. (2017b). Drawing into the AR-Canvas: Designing Embedded Visualizations for Augmented Reality. In *Workshop on Immersive Analytics, IEEE Vis*.

[14] Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., and Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 14(4):379–393, doi: 10.1162/105474605774785235.

[15] Baltabayev, A., Gluschkow, A., Blank, J., Birkh^lzer, G., Kern, M., Klopfer, F., Mayer, L.-M., Scheibler, G., Klein, K., Schreiber, F., and Sommer, B. (2018). Virtual reality for sensor data visualization and analysis. *Electronic Imaging*, 2018(3):451–1.

[16] Barrie, A., Cassell, B., and Cooper, M. (2005). Benefits of Immersion for Viewing 3D Data.

[17] Beetz, M., Hoyningen-huene, N., Kirchlechner, B., Gedikli, S., Siles, F., Durus, M., and Lames, M. (2009). ASPOGAMO : Automated Sports Game Analysis Models. *International Journal of Computer Science in Sport*, 8(1):4–12, doi: 10.1.1.184.9785.

[18] Belcher, D., Billinghurst, M., Hayes, S., and Stiles, R. (2003). Using augmented reality for visualizing complex graphs in three dimensions. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 84–, Washington, DC, USA. IEEE Computer Society.

[19] Bellgardt, M., Gebhardt, S., Hentschel, B., and Kuhlen, T. (2017). Gistualizer: An Immersive Glyph for Multidimensional Datapoints. Workshop on Immersive Analytics, Phoenix (USA).

[20] benchmarks.ul (2021). VR HMD Benchmarks Website benchmarking vr hmds. `https://benchmarks.ul.com/compare/best-vr-headsets`. Accessed: 2021-10-01.

[21] Benford, S., Bowers, J., Fahlen, L. E., Greenhalgh, C., and Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 1, pages 242–249, doi: 10.1145/223904.223935.

[22] Bertin, J. (2010). *Semiology of Graphics: Diagrams, Networks, Maps*. Esri Press, Redlands.

[23] Bhandari, J., MacNeilage, P., and Folmer, E. (2018). Teleportation without spatial disorientation using optical flow cues. In *Proceedings - Graphics Interface*, volume 2018-May, pages 153–158, doi: 10.20380/GI2018.20.

[24] Bialkowski, A., Denman, S., Sridharan, S., Fookes, C., and Lucey, P. (2012). A database for person re-identification in multi-camera surveillance networks. In *2012 International Conference on Digital Image Computing Techniques and Applications, DICTA 2012*, pages 1–8. IEEE, doi: 10.1109/DICTA.2012.6411689.

[25] Bidarra, R., Gambon, D., Kooij, R., Nagel, D., Schutjes, M., and Tziouvara, I. (2013). Gaming at the dentist's – serious game design for pain and discomfort distraction. In *Games for Health*, pages 207–215. Springer, doi: 10.1007/978-3-658-02897-8_16.

[26] Billow, T. V. and Cottam, J. A. (2017). Exploring the Use of Heuristics for Evaluation of an Immersive Analytic System. *Workshop on Immersive Analytics at IEEE Vis*, pages 1–5.

[27] Bliss, J. P., Tidwell, P. D., and Guest, M. A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence: Teleoperators and Virtual Environments*, 6(1):73–86, doi: 10.1162/pres.1997.6.1.73.

[28] Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., Pfister, H., and Oliva, A. (2016). Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, doi: 10.1109/TVCG.2015.2467732.

[29] Borkin, M. A., Gajos, K. Z., Peters, A., Mitsouras, D., Melchionna, S., Rybicki, F. J., Feldman, C. L., and Pfister, H. (2011). Evaluation of artery visualizations for heart disease diagnosis. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2479–2488, doi: 10.1109/TVCG.2011.192.

[30] Bowman, D. A., Davis, E. T., Hodges, L. F., and Badre, A. N. (1999). Maintaining spatial orientation during travel in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 8(6):618–631, doi: 10.1162/105474699566521.

[31] Bowman, D. A. and McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43, doi: 10.1109/MC.2007.257.

[32] Braga, A. A., Papachristos, A. V., and Hureau, D. M. (2014). The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis. *Justice Quarterly*, 31(4):633–663, doi: 10.1080/07418825.2012.673632.

[33] Brase, C. H. and Brase, C. P. (2011). *Understandable statistics: Concepts and methods*. Cengage Learning.

[34] Brath, R. (2014). 3d infovis is here to stay: Deal with it. In *2014 IEEE VIS International Workshop on 3DVis (3DVis)*, pages 25–31, doi: 10.1109/3DVis.2014.7160096.

[35] Brehmer, M. and Munzner, T. (2013). A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, doi: 10.1109/TVCG.2013.124.

[36] Brehmer, M., Ng, J., Tate, K., and Munzner, T. (2016). Matches, Mismatches, and Methods: Multiple-View Workflows for Energy Portfolio Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):449–458, doi: 10.1109/TVCG.2015.2466971.

[37] Brodbeck, D. and Girardin, L. (2003). Design study: Using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In *Proceedings - International Conference on Coordinated and Multiple Views in Exploratory Visualization, CMV 2003, Associated and co-located with International Conference on Information Visualization, IV 2003*, pages 104–111, London UK. IEEE, CMV, doi: 10.1109/CMV.2003.1215008.

[38] Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7.

[39] Brooks, F. P. (1999). What's real about virtual reality? *IEEE Computer Graphics Application*, 19(6):16–27, doi: 10.1109/38.799723.

[40] Brooks, J., Nagels, S., and Lopes, P. (2020). Trigeminal-based Temperature Illusions. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–12, doi: 10.1145/3313831.3376806.

[41] Brown, H. A. and Marnett, L. J. (2011). Introduction to lipid biochemistry, metabolism, and signaling. In *Chemical Reviews*, volume 111 of *Emerging communication : studies in new technologies and practices in communication*, pages 5817–5820. IOS Press, Netherlands, doi: 10.1021/cr200363s.

[42] Bryson, S. (1996). Virtual reality in scientific visualization. *Communications of the ACM*, 39(5):62–71.

[43] Buchmüller, J., Jäckle, D., Cakmak, E., Brandes, U., and Keim, D. A. (2019). MotionRugs: Visualizing Collective Trends in Space and Time. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):76–86, doi: 10.1109/TVCG.2018.2865049.

[44] Burigat, S. and Chittaro, L. (2007). Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. *International Journal of Human Computer Studies*, 65(11):945–958, doi: 10.1016/j.ijhcs.2007.07.003.

[45] Büring, T., Gerken, J., and Reiterer, H. (2006). Usability of overview-supported zooming on small screens with regard to individual differences in spatial ability. In *Proceedings of the Workshop on Advanced Visual Interfaces*, volume 2006, pages 233–240, doi: 10.1145/1133265.1133310.

[46] Büschel, W., Reipschläger, P., Langner, R., and Dachselt, R. (2017). Investigating the use of spatial interaction for 3D data visualization on mobile devices. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, ISS 2017*, pages 62–71, New York, NY, USA. ACM, doi: 10.1145/3132272.3134125.

[47] Butcher, P. W. S., John, N. W., and Ritsos, P. D. (2020). VRIA: A Web-based Framework for Creating Immersive Analytics Experiences. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, doi: 10.1109/tvcg.2020.2965109.

[48] Butscher, S., Hubenschmid, S., Müller, J., Fuchs, J., and Reiterer, H. (2018). Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 90. ACM.

[49] Byška, J., Le Muzic, M., Gröller, M. E., Viola, I., and Kozlíková, B. (2016). AnimoAminoMiner: Exploration of Protein Tunnels and their Properties in Molecular Dynamics. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):747–756, doi: 10.1109/TVCG.2015.2467434.

[50] Cánovas, R., León, I., Roldán, M. D., Astur, R., and Cimadevilla, J. M. (2009). Virtual reality tasks disclose spatial memory alterations in fibromyalgia. *Rheumatology (Oxford, England)*, 48(10):1273–1278, doi: 10.1093/rheumatology/kep218.

[51] Cao, N., Shi, C., Lin, S., Lu, J., Lin, Y. R., and Lin, C. Y. (2016). TargetVue: Visual Analysis of Anomalous User Behaviors in Online Communication Systems. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):280–289, doi: 10.1109/TVCG.2015.2467196.

[52] Carpendale, S. (2008). Evaluating information visualizations. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4950 LNCS, pages 19–45. Springer, doi: 10.1007/978-3-540-70956-5_2.

[53] Carr, D. B., Littlefield, R. J., Nicholson, W. L., and Littlefield, J. S. (1987). Scatterplot matrix techniques for large N. *Journal of the American Statistical Association*, 82(398):424–436, doi: 10.1080/01621459.1987.10478445.

[54] Cavallo, M. and Demiralp, Ç. (2019). Clustrophile 2: Guided Visual Clustering Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):267–276, doi: 10.1109/TVCG.2018.2864477.

[55] Cernea, D. (2020). Multi-View Stereo Reconstruction Library. `https://cdcseacave.github.io/openMVS`. Accessed: 2018-08-17.

[56] Chandler, T., Cordeil, M., Czauderna, T., Dwyer, T., Glowacki, J., Goncu, C., Klapperstueck, M., Klein, K., Marriott, K., Schreiber, F., and Wilson, E. (2015). Immersive Analytics. In *2015 Big Data Visual Analytics, BDVA 2015*, pages 1–8. IEEE, doi: 10.1109/BDVA.2015.7314296.

[57] Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *AMC Computing Surveys (CSUR)*, 41(3):15.

[58] Changsong, S., Chris, Z., and Fels, S. (2006). A multi-camera surveillance system that estimates quality-of-view measurement. In *Proceedings - International Conference on Image Processing, ICIP*, volume 3, pages III—-193. IEEE, doi: 10.1109/ICIP.2007.4379279.

[59] Checa, D. and Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications*, 79(9):5501–5527.

[60] Chen, M., Gaither, K., John, N. W., and McCann, B. (2019). An Information-Theoretic Approach to the Cost-benefit Analysis of Visualization in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):32–42, doi: 10.1109/TVCG.2018.2865025.

[61] Chen, M. and Golan, A. (2016). What May Visualization Processes Optimize? *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2619–2632, doi: 10.1109/TVCG.2015.2513410.

[62] Chi, E. H. (2000). Taxonomy of visualization techniques using the data state reference model. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 69–75, doi: 10.1109/infvis.2000.885092.

[63] Chittaro, L. and Burigat, S. (2004). 3D Location-pointing as a navigation aid in virtual environments. In *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, pages 267–274, doi: 10.1145/989863.989910.

[64] Clark, H. H. and Brennan, S. E. (2004). Grounding in communication. *Perspectives on socially shared cognition.*, pages 127–149, doi: 10.1037/10096-006.

[65] Cleveland, W. S. and McGill, R. (1984). The many faces of a scatterplot. *Journal of the American Statistical Association*, 79(388):807–822, doi: 10.1080/01621459.1984.10477098.

[66] Cliburn, D., Winlock, T., Rilea, S., and Van Donsel, M. (2007). Dynamic landmark placement as a navigation aid in virtual worlds. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, pages 211–214. ACM, doi: 10.1145/1315184.1315225.

[67] Cockburn, A. and McKenzie, B. (2002). Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 4, pages 203–210, doi: 10.1145/503376.503413.

[68] Cohen, S., Glaser, B. G., and Strauss, A. L. (1969). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, volume 20. Aldine, Chicago, doi: 10.2307/588533.

[69] Collins, R. T., Lipton, A. J., Fujiyoshi, H., and Kanade, T. (2001). Algorithms for cooperative multisensor surveillance. *Proceedings of the IEEE*, 89(10):1456–1477, doi: 10.1109/5.959341.

[70] Coluccia, E. and Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology*, 24(3):329–340, doi: 10.1016/j.jenvp.2004.08.006.

[71] Cordeil, M., Bach, B., Cunningham, A., Montoya, B., Smith, R. T., Thomas, B. H., and Dwyer, T. (2020). Embodied Axes: Tangible, Actuated Interaction for 3D Augmented Reality Data Spaces. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–12, doi: 10.1145/3313831.3376613.

[72] Cordeil, M., Bach, B., Li, Y., Wilson, E., and Dwyer, T. (2017a). Design space for spatio-data coordination: Tangible interaction devices for immersive information visualisation. In *IEEE Pacific Visualization Symposium*, pages 46–50, doi: 10.1109/PACIFICVIS.2017.8031578.

[73] Cordeil, M., Cunningham, A., Bach, B., Hurter, C., Thomas, B. H., Marriott, K., and Dwyer, T. (2019). Iatk: An immersive analytics toolkit. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 200–209. IEEE.

[74] Cordeil, M., Cunningham, A., Dwyer, T., Thomas, B. H., and Marriott, K. (2017b). 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*, pages 71–83. ACM.

[75] Cordeil, M., Dwyer, T., Klein, K., Laha, B., Marriott, K., and Thomas, B. H. (2017c). Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on Visualization and Computer Graphics*, 23(1):441–450, doi: 10.1109/TVCG.2016.2599107.

[76] Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1993*, pages 135–142. ACM, doi: 10.1145/166117.166134.

[77] Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. 35(6):64–72.

[78] Cummings, J. J. and Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, 19(2):272–309, doi: 10.1080/15213269.2015.1015740.

[79] Czauderna, T., Haga, J., Kim, J., Klapperstueck, M., Klein, K., Kuhlen, T. W., Oeltze-Jafra, S., Sommer, B., and Schreiber, F. (2018). *Immersive Analytics Applications in Life and Health Sciences*, pages 289–330. Springer International Publishing.

[80] Czerwinski, M. P., van Dantzich, M., Robertson, G., and Hoffman, H. (1999). The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D. In *Press*, pages 163–170, doi: 10.1.1.33.774.

[81] Dai, J., Li, Y., He, K., and Sun, J. (2016). R-FCN: Object detection via region-based fully convolutional networks. In *Advances in Neural Information Processing Systems*, pages 379–387.

[82] Darken, R. and Peterson, B. (2014). Spatial Orientation, Wayfinding, and Representation. *Handbk Virtual Environ*, 2002:467–491, doi: 10.1201/b17360-24.

[83] Darken, R. P. and Banker, W. P. (1998). Navigating in natural environments: A virtual environment training transfer study. In *Proceedings - Virtual Reality Annual International Symposium*, pages 12–19. IEEE, IEEE, doi: 10.1109/vrais.1998.658417.

[84] Darken, R. P. and Cevik, H. (1999). Map usage in virtual environments: orientation issues. In *Proceedings - Virtual Reality Annual International Symposium*, pages 133–140. IEEE, IEEE, doi: 10.1109/vr.1999.756944.

[85] Darken, R. P., Cockayne, W. R., and Carmein, D. (1997). Omni-directional treadmill: A locomotion device for virtual worlds. In *UIST (User Interface Software and Technology): Proceedings of the ACM Symposium*, pages 213–221. ACM.

[86] Darken, R. P. and Sibert, J. L. (1993). A toolset for navigation in virtual environments. In *UIST 1993 - Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, pages 157–165. ACM, doi: 10.1145/168642.168658.

[87] Darken, R. P. and Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 142–149. ACM, doi: 10.1145/238386.238459.

[88] Dasgupta, D. and Forrest, S. (1999). Novelty detection in time series data using ideas from immunology. In *The 8th International Conference on Intelligent Systems*, pages 82–87.

[89] Davies, J. H. (2013). Global map of solid Earth surface heat flow. *Geochemistry, Geophysics, Geosystems*, 14(10):4608–4622, doi: 10.1002/ggge.20271.

[90] Dehingia, M., Devi, K. T., Talukdar, N. C., Talukdar, R., Reddy, N., Mande, S. S., Deka, M., and Khan, M. R. (2015). Gut bacterial diversity of the tribes of India and comparison with the worldwide data. *Scientific Reports*, 5(18563):2045–2322, doi: 10.1038/srep18563.

[91] Devanne, M., Wannous, H., Berretti, S., Pala, P., Daoudi, M., and Del Bimbo, A. (2015). 3-D Human Action Recognition by Shape Analysis of Motion Trajectories on Riemannian Manifold. *IEEE Transactions on Cybernetics*, 45(7):1340–1352, doi: 10.1109/TCYB.2014.2350774.

[92] Dictionary (2018). Dictionary Metaphor. `https://www.merriam-webster.com/dictionary/metaphor`. Accessed: 2018-07-19.

[93] Dictionary (2019a). Reality: Dictionary. `http://www.dictionary.com/browse/reality`. Accessed: 25-May-2019.

[94] Dictionary (2019b). Virtual: Dictionary. `http://www.dictionary.com/browse/virtual`. Accessed: 25-May-2019.

[95] Dodge, M., Mcderby, M., and Turner, M. (2008). *Geographic Visualization: Concepts, Tools and Applications*, doi: 10.1002/9780470987643.

[96] Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., Yeh, S., Mahabal, A., Graham, M. J., Drake, A. J., Davidoff, S., Norris, J. S., Longo, G., Cioc, A., Wang, A., Longo, G., Norris, J. S., Zhang, J., Lawler, E., Yeh, S., Mahabal, A., Graham, M. J., and Drake, A. J. (2014). Immersive and Collaborative Data Visualization Using Virtual Reality Platforms. *CoRR*, abs/1410.7670:609–614, doi: 10.1109/BigData.2014.7004282.

[97] Dos Santos, S. and Brodlie, K. (2004). Gaining understanding of multivariate and multidimensional data through visualization. *Computers and Graphics (Pergamon)*, 28(3):311–325, doi: 10.1016/j.cag.2004.03.013.

[98] Drogemuller, A., Cunningham, A., Walsh, J., Cordeil, M., Ross, W., and Thomas, B. (2018). Evaluating navigation techniques for 3d graph visualizations in virtual reality. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–10. IEEE.

[99] Duncan, D., Garner, R., Zrantchev, I., Ard, T., Newman, B., Saslow, A., Wanserski, E., and Toga, A. W. (2019). Using Virtual Reality to Improve Performance and User Experience in Manual Correction of MRI Segmentation Errors by Non-experts. *Journal of Digital Imaging*, 32(1):97–104, doi: 10.1007/s10278-018-0108-5.

[100] Dünser, A., Steinbügl, K., Kaufmann, H., and Glück, J. (2006). Virtual and augmented reality as spatial ability training tools. In *ACM International Conference Proceeding Series*, volume 158, pages 125–132. ACM, doi: 10.1145/1152760.1152776.

[101] Dutta, S., Chen, C. M., Heinlein, G., Shen, H. W., and Chen, J. P. (2017). In Situ Distribution Guided Analysis and Visualization of Transonic Jet Engine Simulations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):811–820, doi: 10.1109/TVCG.2016.2598604.

[102] Dwyer, T., Marriott, K., Isenberg, T., Klein, K., Riche, N., Schreiber, F., Stuerzlinger, W., and Thomas, B. H. (2018). Immersive analytics: an introduction. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11190 LNCS, pages 1–23. Springer, 6330 Cham, Switzerland, doi: 10.1007/978-3-030-01388-2_1.

[103] Eick, S. G. (1995). Engineering perceptually effective visualizations for abstract data. In *In Scientific Visualization Overviews, Methodologies and Techniques, IEEE Computer Science*. Citeseer.

[104] Ellis, S. R. (1994). What Are Virtual Environments? *IEEE Computer Graphics and Applications*, 14(1):17–22, doi: 10.1109/38.250914.

[105] Elmqvis, N. (2018). Blog Entry - 3D Visualization for Nonspatial Data: Guidelines and Challenges. https://sites.umiacs.umd.edu/elm/2017/10/03/3d-visualization-for-nonspatial-data-guidelines-and-challenges/. Accessed: 2018-07-19.

[106] Elmqvist, N., Dragicevic, P., and Fekete, J. D. (2008). Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1141–1148, doi: 10.1109/TVCG.2008.153.

[107] ElSayed, N. A., Thomas, B. H., Marriott, K., Piantadosi, J., and Smith, R. T. (2016). Situated Analytics: Demonstrating immersive analytical tools with Augmented Reality. *Journal of Visual Languages and Computing*, 36:13–23, doi: 10.1016/j.jvlc.2016.07.006.

[108] Ens, B., Bach, B., Cordeil, M., Engelke, U., Serrano, M., Willett, W., Prouzeau, A., Anthes, C., Büschel, W., Dunne, C., et al. (2021). Grand challenges in immersive analytics. In *CHI Conference on Human Factors in Computing Systems (CHI'21)*, pages 1–17. ACM.

[109] Erra, U., Malandrino, D., and Pepe, L. (2018). Virtual reality interfaces for interacting with three-dimensional graphs. *International Journal of Human–Computer Interaction*, 35(1):1–14, doi: 10.1080/10447318.2018.1429061.

[110] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pages 226–231.

[111] Etemadpour, R., Monson, E., and Linsen, L. (2013). The effect of stereoscopic immersive environments on projection-based multi-dimensional data visualization. In *Proceedings of the International Conference on Information Visualisation*, pages 389–397, doi: 10.1109/IV.2013.51.

[112] Etemadpour, R., Olk, B., and Linsen, L. (2014). Eye-tracking investigation during visual analysis of projected multidimensional data with 2D scatterplots. In *IVAPP 2014 - Proceedings of the 5th International Conference on Information Visualization Theory and Applications*, pages 233–246. IEEE, doi: 10.5220/0004675802330246.

[113] Faust, R., Glickenstein, D., and Scheidegger, C. (2019). DimReader: Axis lines that explain non-linear projections. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):481–490, doi: 10.1109/TVCG.2018.2865194.

[114] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–53.

[115] Férey, N., Gros, P.-E., Hérisson, J., and Gherbi, R. (2005). Visual data mining of genomic databases by immersive graph-based exploration. In *Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 143–146.

[116] Filho, J. A. W., Rey, M. F., Freitas, C. M., and Nedel, L. (2017). Immersive analytics of dimensionality-reduced data scatterplots. In *Proceedings of the Workshop on Immersive Analytics*. Workshop on Immersive Analytics, Phoenix (USA).

[117] Fisher, D. (2007). Hotmap: Looking at geographic attention. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1184–1191, doi: 10.1109/TVCG.2007.70561.

[118] FLORIDA (2020). FLORIDA - Flexibles, teilautomatisiertes Analysesystem zur Auswertung von Videomassendaten. `http://www.florida-project.de/`. Accessed: 25-May-2019.

[119] Fluke, C. J. and Barnes, D. G. (2016). The Ultimate Display. *Multimedia: From Wagner to virtual reality*, pages 506–508.

[120] Fonnet, A. and Prie, Y. (2019). Survey of Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, doi: 10.1109/tvcg.2019.2929033.

[121] Fonnet, A., Vigier, T., Prie, Y., Cliquet, G., and Picarougne, F. (2018). Axes and coordinate systems representations for immersive analytics of multi-dimensional data. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–10. IEEE.

[122] Forsberg, A. S., Laidlaw, D. H., Van Dam, A., Kirby, R. M., Karniadakis, G. E., and Elion, J. L. (2000). Immersive virtual reality for visualizing flow through an artery. In *Proceedings of the IEEE Visualization Conference*, pages 457–460+589, Los Alamitos, CA, USA. IEEE, doi: 10.1109/visual.2000.885731.

[123] Freksa, C. (2004). Spatial cognition: An AI perspective. In *Frontiers in Artificial Intelligence and Applications*, volume 110, pages 1122–1128. IOS, Press, doi: 10.5555/3000001.3000289.

[124] FTW (2018). Design Principles FTW - Desiging for VR. `https://www.designprinciplesftw.com/collections/designing-for-virtual-reality`. Accessed: 2018-07-19.

[125] Fuchs, J., Isenberg, P., Bezerianos, A., and Keim, D. (2017). A systematic review of experimental studies on data glyphs. *IEEE Transactions on Visualization and Computer Graphics*, 23(7):1863–1879, doi: 10.1109/TVCG.2016.2549018.

[126] Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., and Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 529–536. ACM, doi: 10.1145/642611.642703.

[127] García-Alonso, C. R., Pérez-Naranjo, L. M., and Fernández-Caballero, J. C. (2014). Multiobjective evolutionary algorithms to identify highly autocorrelated areas: The case of spatial distribution in financially compromised farms. *Annals of Operations Research*, 219(1):187–202, doi: 10.1007/s10479-011-0841-3.

[128] García-Hernández, R. J., Anthes, C., Wiedemann, M., and Kranzlmüller, D. (2016). Perspectives for using virtual reality to extend visual data mining in information visualization. In *2016 IEEE Aerospace Conference*, pages 1–11, doi: 10.1109/AERO.2016.7500608.

[129] Gatalsky, P., Andrienko, N., and Andrienko, G. (2004). Interactive analysis of event data using space-time cube. In *Proceedings of the International Conference on Information Visualization*, volume 8, pages 145–152, doi: 10.1109/iv.2004.1320137.

[130] Gerardin, P., Kourtzi, Z., and Mamassian, P. (2010). Prior knowledge of illumination for 3D perception in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(37):16309–16314, doi: 10.1073/pnas.1006285107.

[131] Gershon, N. (1998). Visualization of an imperfect world. *IEEE Computer Graphics and Applications*, 18(4):43–45, doi: 10.1109/38.689662.

[132] Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C. D., and Roberts, J. C. (2011). Visual comparison for information visualization. *Information Visualization*, 10(4):289–309, doi: 10.1177/1473871611416549.

[133] Goffredo, M., Bouchrika, I., Carter, J. N., and Nixon, M. S. (2010). Performance analysis for automated gait extraction and recognition in multi-camera surveillance. *Multimedia Tools and Applications*, 50(1):75–94, doi: 10.1007/s11042-009-0378-5.

[134] Google (2018). Google Carboard Design Guidelines VR. `https://designguidelines.withgoogle.com/cardboard/designing-for-google-cardboard`. Accessed: 2018-07-19.

[135] Google (2019). Google Earth Computer Program. `https://www.google.com/earth/`. Accessed: 2019-09-17.

[136] Google (2020). Google VR: Cardboard. `https://www.google.com/get/cardboard`. Accessed: 14-July-2019.

[137] Gracia, A., González, S., Robles, V., Menasalvas, E., and Von Landesberger, T. (2016). New insights into the suitability of the third dimension for visualizing multivariate/multidimensional data: A study based on loss of quality quantification. *Information Visualization*, 15(1):3–30, doi: 10.1177/1473871614556393.

[138] Grammenos, D., Filou, M., Papadakos, P., and Stephanidis, C. (2002). Virtual Prints : Leaving trails in Virtual Environments. In *EGVE '02: Proceedings of the workshop on Virtual environments 2002*, volume 2, pages 131–139. The Eurographics Association, doi: 10.2312/EGVE/EGVE02/131-138.

[139] Greffard, N., Picarougne, F., and Kuntz, P. (2014). Beyond the classical monoscopic 3d in graph analytics: an experimental study of the impact of stereoscopy. In *3DVis (3DVis), 2014 IEEE VIS International Workshop on*, pages 19–24. IEEE.

[140] Grossman, T. and Balakrishnan, R. (2006). An evaluation of depth perception on volumetric displays. In *Proceedings of the Workshop on Advanced Visual Interfaces*, volume 2006 of *AVI '06*, pages 193–200, New York, NY, USA. ACM, doi: 10.1145/1133265.1133305.

[141] Gruchalla, K. (2004). Immersive well-path editing: Investigating the added value of immersion. In *Proceedings - Virtual Reality Annual International Symposium*, pages 157–164, doi: 10.1109/VR.2004.1310069.

[142] Günther, T. and Theisel, H. (2017). Backward Finite-Time Lyapunov Exponents in Inertial Flows. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):970–979, doi: 10.1109/TVCG.2016.2599016.

[143] Günther, U., Pietzsch, T., Gupta, A., Harrington, K. I., Tomancak, P., Gumhold, S., and Sbalzarini, I. F. (2019). scenery: Flexible virtual reality visualization on the java vm. In *2019 IEEE Visualization Conference (VIS)*, pages 1–5. IEEE.

[144] Gustafson, S. G. and Irani, P. P. (2007). Comparing visualizations for tracking off-screen moving targets. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 2399–2404, doi: 10.1145/1240866.1241014.

[145] Gutiérrez, F., Pierce, J., Vergara, V., Coulter, R., Saland, L., Caudell, T., Goldsmith, T., and Alverson, D. (2007). the Impact of the Degree of Immersion Upon Learning Performance in Virtual Reality Simulations for Medical Education. In *Journal of Investigative Medicine*, volume 55, page S91, doi: 10.1097/00042871-200701010-00099.

[146] Haber, R. B. and Mcnabb, D. A. (1990). Visualization ldioms: A Conceptual Model for Scientific Visualization Systems. *Visualization in scientific computing*, 74:93.

[147] Haddad, R. A. and Akansu, A. N. (1991). A Class of Fast Gaussian Binomial Filters for Speech and Image Processing. *IEEE Transactions on Signal Processing*, 39(3):723–727, doi: 10.1109/78.80892.

[148] Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, Amsterdam, Netherlands, doi: 10.1016/S0166-4115(08)62386-9.

[149] Hayatpur, D., Xia, H., and Wigdor, D. (2020). DataHop: Spatial data exploration in virtual reality. In *UIST 2020 - Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pages 818–828, doi: 10.1145/3379337.3415878.

[150] He, K., Zhang, X., Ren, S., and Sun, J. (2014). Spatial pyramid pooling in deep convolutional networks for visual recognition. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 8691 LNCS, pages 346–361. Springer, doi: 10.1007/978-3-319-10578-9_23.

[151] Hegarty, M. and Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2):175–191, doi: 10.1016/j.intell.2003.12.001.

[152] Helbig, C., Bauer, H.-S., Rink, K., Wulfmeyer, V., Frank, M., and Kolditz, O. (2014). Concept and workflow for 3d visualization of atmospheric data in a virtual reality environment for analytical approaches. *Environmental Earth Sciences*, 72(10):3767–3780, doi: 10.1007/s12665-014-3136-6.

[153] Hendrix, C. and Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments*, 5(3):274–289, doi: 10.1162/pres.1996.5.3.274.

[154] Hettinger, L. J. and Riccio, G. E. (1992). Visually Induced Motion Sickness in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 1(3):306–310, doi: 10.1162/pres.1992.1.3.306.

[155] Hibbard, P. B., Haines, A. E., and Hornsey, R. L. (2017). Magnitude, precision, and realism of depth perception in stereoscopic vision. *Cognitive Research: Principles and Implications*, 2(1):25, doi: 10.1186/s41235-017-0062-7.

[156] Hodge, V. and Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126, doi: 10.1007/s10462-004-4304-y.

[157] Holten, D. and Van Wijk, J. J. (2010). Evaluation of cluster identification performance for different PCP variants. *Computer Graphics Forum*, 29(3):793–802, doi: 10.1111/j.1467-8659.2009.01666.x.

[158] Homps, F., Beugin, Y., and Vuillemot, R. (2020). ReViVD: Exploration and Filtering of Trajectories in an Immersive Environment using 3D Shapes. In *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*, pages 729–737. IEEE, doi: 10.1109/VR46266.2020.1581269207852.

[159] Hsieh, T.-J., Olsen, M. J., Johnstone, E., Young, A. P., Driscoll, N., Ashford, S. A., and Kuester, F. (2007). VR-based visual analytics of LIDAR data for cliff erosion assessment. In *Proc. ACM Symposium on Virtual Reality Software and Technology*, VRST '07, pages 249–250. ACM, doi: 10.1145/1315184.1315244.

[160] HTC (2017). VIVE | Controller. `https://vive.com/eu/accessory/controller/`. Accessed: 25-May-2019.

[161] HTC (2018). HTC Vive Developer Website. `https://www.vive.com/eu/`. Accessed: 2018-08-17.

[162] Huang, B., Jiang, B., and Li, H. (2001). An integration of GIS, virtual reality and the internet for visualization, analysis and exploration of spatial data. *International Journal of Geographical Information Science*, 15(5):439–456, doi: 10.1080/13658810110046574.

[163] Huang, Q., Mo, W., Pei, L., and Zeng, L. (2019). 3D coordinate calculation and pose estimation of power meter based on binocular stereo vision. In *ACM International Conference Proceeding Series*, volume Part F147956, pages 282–285. ACM, doi: 10.1145/3316615.3316655.

[164] Hubmann-Haidvogel, A., Scharl, A., and Weichselbraun, A. (2009). Multiple coordinated views for searching and navigating Web content repositories. *Information Sciences*, 179(12):1813–1821, doi: 10.1016/j.ins.2009.01.030.

[165] Hurley, C. B. (2004). Clustering visualizations of multidimensional data. *Journal of Computational and Graphical Statistics*, 13(4):788–806, doi: 10.1198/106186004X12425.

[166] Hurter, C., Riche, N. H., Drucker, S. M., Cordeil, M., Alligier, R., and Vuillemot, R. (2019). Fiberclay: Sculpting three dimensional trajectories to reveal structural insights. *IEEE transactions on visualization and computer graphics*, 25(1):704–714.

[167] Hvass, J., Larsen, O., Vendelbo, K., Nilsson, N., Nordahl, R., and Serafin, S. (2018). Visual realism and presence in a virtual reality game. In *3DTV-Conference*, volume 2017-June, pages 1–4. IEEE, doi: 10.1109/3DTV.2017.8280421.

[168] Hyde, D. A., Hall, T. R., and Caers, J. (2018). Vrge: An immersive visualization application for the geosciences. In *2018 IEEE Scientific Visualization Conference (SciVis)*, pages 1–5. IEEE.

[169] Inselberg, A. and Dimsdale, B. (1987). Parallel Coordinates for Visualizing Multi-Dimensional Geometry BT - Computer Graphics 1987. In *Computer Graphics 1987*, pages 25–44. Springer, doi: 10.1007/978-4-431-68057-4_3.

[170] Interrante, V., Fuchs, H., and Pizer, S. M. (1997). Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3(2):98–117, doi: 10.1109/2945.597794.

[171] Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M., and Moller, T. (2013). A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, doi: 10.1109/TVCG.2013.126.

[172] ISPR (2018). Intl. Society for Presence Research - Presence. `https://ispr.info/about-presence-2/about-presence/`. Accessed: 2018-07-21.

[173] Jaeger, S., Klein, K., Joos, L., Zagermann, J., De Ridder, M., Kim, J., Yang, J., Pfeil, U., Reiterer, H., and Schreiber, F. (2019). Challenges for brain data analysis in VR environments. In *IEEE Pacific Visualization Symposium*, volume 2019-April, pages 42–46, doi: 10.1109/PacificVis.2019.00013.

[174] Janetzko, H., Hao, M. C., Mittelstädt, S., Dayal, U., and Keim, D. (2013). Enhancing scatter plots using ellipsoid pixel placement and shading. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, pages 1522–1531. IEEE, doi: 10.1109/HICSS.2013.197.

[175] Jansen, Y., Dragicevic, P., and Fekete, J. D. (2013). Evaluating the efficiency of physical visualizations. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 2593–2602. ACM, doi: 10.1145/2470654.2481359.

[176] Javed, O., Rasheed, Z., Alatas, O., and Shah, M. (2003). KNIGHT™: A real time surveillance system for multiple and non-overlapping cameras. In *Proceedings - IEEE International Conference on Multimedia and Expo*, volume 1, pages I649–I652. IEEE, doi: 10.1109/ICME.2003.1221001.

[177] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human Computer Studies*, 66(9):641–661, doi: 10.1016/j.ijhcs.2008.04.004.

[178] Jentner, W., Sacha, D., Stoffel, F., Ellis, G., Zhang, L., and Keim, D. A. (2018). Making machine intelligence less scary for criminal analysts: reflections on designing a visual comparative case analysis tool. *Visual Computer*, 34(9):1225–1241, doi: 10.1007/s00371-018-1483-0.

[179] Jeong, C. S. and Pang, A. (1998). Reconfigurable disc trees for visualizing large hierarchical information space. In Wills, G. and Dill, J. (Eds.), *Proceedings of the IEEE Symposium on Information Visualization*, pages 19–25. IEEE Computer Society, doi: 10.1109/infvis.1998.729555.

[180] Jo, J., Vernier, F., Dragicevic, P., and Fekete, J.-D. D. (2019). A Declarative Rendering Model for Multiclass Density Maps. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):470–480, doi: 10.1109/TVCG.2018.2865141.

[181] Kageyama, A., Tamura, Y., and Sato, T. (2000). Visualization of vector field by virtual reality. *Progress of Theoretical Physics Supplement*, 138:665–673, doi: 10.1143/PTPS.138.665.

[182] Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., and Melançon, G. (2008a). Visual analytics: Definition, process, and challenges. In *Information Visualization*, pages 154–175. Springer.

[183] Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., and Ziegler, H. (2008b). Visual analytics: Scope and challenges. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4404 LNCS, pages 76–90. Springer, doi: 10.1007/978-3-540-71080-6_6.

[184] Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, doi: 10.1207/s15327108ijap0303_3.

[185] Kerren, A. and Schreiber, F. (2014). Why integrate infovis and scivis? an example from systems biology. *IEEE Computer Graphics and Applications*, 34(6):69–73.

[186] Kettnaker, V. and Zabih, R. (1999). Bayesian multi-camera surveillance. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 253–259. IEEE, doi: 10.1109/cvpr.1999.784638.

[187] Kim, K., Maloney, D., Bruder, G., Bailenson, J. N., and Welch, G. F. (2017). The effects of virtual human's spatial and behavioral coherence with physical objects on social presence in AR. *Computer Animation and Virtual Worlds*, 28(3-4):e1771, doi: 10.1002/cav.1771.

[188] Kindt, E. J. (2018). Having yes, using no? About the new legal regime for biometric data. *Computer Law and Security Review*, 34(3):523–538, doi: 10.1016/j.clsr.2017.11.004.

[189] Klemm, P., Lawonn, K., Glaßer, S., Niemann, U., Hegenscheid, K., Völzke, H., and Preim, B. (2016). 3D Regression Heat Map Analysis of Population Study Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):81–90, doi: 10.1109/TVCG.2015.2468291.

[190] Kohlhammer, J., Keim, D., Pohl, M., Santucci, G., and Andrienko, G. (2011). *Solving problems with visual analytics*, volume 7. Eurographics Association, doi: 10.1016/j.procs.2011.12.035.

[191] König, S., Keshava, A., Clay, V., Rittershofer, K., Kuske, N., and König, P. (2020). Embodied Spatial Knowledge Acquisition in Immersive Virtual Reality: Comparison of Direct Experience and Map Exploration. *bioRxiv*, doi: 10.1101/2020.01.12.903096.

[192] Koren, Y. and Carmel, L. (2004). Robust linear dimensionality reduction. *IEEE Transactions on Visualization and Computer Graphics*, 10(4):459–470, doi: 10.1109/TVCG.2004.17.

[193] Kosara, R., Sahling, G. N., and Hauser, H. (2004). Linking Scientific and Information Visualization with Interactive 3D Scatterplots. *Short Communication Papers Proceedings of the 12th International Conference in Central Europe on Computer Graphics, Visualization, and Computer Vision (WSCG)*, pages 133–140.

[194] Kotlarek, J., Lin, I. C., and Ma, K. L. (2018). Improving spatial orientation in immersive environments. In *SUI 2018 - Proceedings of the Symposium on Spatial User Interaction*, pages 79–88, doi: 10.1145/3267782.3267792.

[195] Kozlíková, B., Krone, M., Falk, M., Lindow, N., Baaden, M., Baum, D., Viola, I., Parulek, J., and Hege, H. C. (2017). Visualization of Biomolecular Structures: State of the Art Revisited. *Computer Graphics Forum*, 36(8):178–204, doi: 10.1111/cgf.13072.

[196] Kraak, M.-j. (2003). The Space-Time Cube Revisited from a Geovisualization Perspective. In *21st International Cartographic Conference (ICC)*, pages 10–16.

[197] Krakov, D. and Feitelson, D. G. (2013). Comparing performance heatmaps. In *Workshop on Job Scheduling Strategies for Parallel Processing*, volume 8429 LNCS, pages 42–61, Berlin, Heidelberg. Springer, doi: 10.1007/978-3-662-43779-7_3.

[198] Kraus, M., Angerbauer, K., Buchmüller, J., Schweitzer, D., Keim, D. A., Sedlmair, M., and Fuchs, J. (2020a). Assessing 2D and 3D Heatmaps for Comparative Analysis: An Empirical Study. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–14, doi: 10.1145/3313831.3376675.

[199] Kraus, M., Buchmüller, J., Schweitzer, D., Keim, D. A., and Fuchs, J. (2019a). Comparative analysis with heightmaps in virtual reality environments. In *EUROVIS 2019: 21st EG/VGTC Conference on Visualization*.

[200] Kraus, M., Fuchs, J., Sommer, B., Klein, K., Engelke, U., Keim, D. A., and Schreiber, F. (2021a). Immersive Analytics with Abstract 3D Visualizations: a Survey. *Computer Graphics Forum*, doi: 10.1111/cgf.14430.

[201] Kraus, M., Klein, K., Fuchs, J., Keim, D. A., Schreiber, F., and Sedlmair, M. (2021b). The value of immersive visualization. *IEEE Computer Graphics and Applications*, 41(4):125–132, doi: 10.1109/MCG.2021.3075258.

[202] Kraus, M., Klein, K., Fuchs, J., Keim, D. A., Schreiber, F., and Sedlmair, M. (2021c). The value of immersive visualization. *IEEE computer graphics and applications*, 41(4):125–132.

[203] Kraus, M., Miller, M., Buchmüller, J., Stein, M., Weiler, N., Keim, D. A., and El-Assady, M. (2020b). Breaking the Curse of Visual Analytics: Accommodating Virtual Reality in the Visualization Pipeline. In Cláudio, A. P., Bouatouch, K., Chessa, M., Paljic, A., Kerren, A., Hurter, C., Tremeau, A., and Farinella, G. M. (Eds.), *Communications in Computer and Information Science*, volume 1182 CCIS, pages 253–284. Springer International Publishing, Cham, doi: 10.1007/978-3-030-41590-7_11.

[204] Kraus, M., Pollok, T., Miller, M., Kilian, T., Moritz, T., Schweitzer, D., Beyerer, J., Keim, D., Qu, C., and Jentner, W. (2020c). Toward mass video data analysis: Interactive and immersive 4d scene reconstruction. *Sensors (Switzerland)*, 20(18):1–38, doi: 10.3390/s20185426.

[205] Kraus, M., Schäfer, H., Meschenmoser, P., Schweitzer, D., Keim, D. A., Sedlmair, M., and Fuchs, J. (2020d). A comparative study of orientation support tools in virtual reality environments with virtual teleportation. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 227–238. IEEE, doi: 10.1109/ISMAR50242.2020.00046.

[206] Kraus, M., Weiler, N., Breitkreutz, T., Keim, D. A., and Stein, M. (2019b). Breaking the curse of visual data exploration: Improving analyses by building bridges between data world and real world. In *VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, volume 3, pages 19–27. INSTICC, SciTePress, doi: 10.5220/0007257400190027.

[207] Kraus, M., Weiler, N., Keim, D. A., Diehl, A., and Bach, B. (2018). Visualization in the VR-Canvas: How much Reality is Good for Immersive Analytics in Virtual Reality? In *VisGuides: 2nd Workshop on the Creation, Curation, Critique and Conditioning of Principles and Guidelines in Visualization*. Retrieved from https://c4pgv.dbvis.de/Kraus_et_al_2018.pdf.

[208] Kraus, M., Weiler, N., Oelke, D., Kehrer, J., Keim, D. A., and Fuchs, J. (2019c). The Impact of Immersion on Cluster Identification Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):525–535, doi: 10.1109/TVCG.2019.2934395.

[209] Krekhov, A., Cmentowski, S., Waschk, A., and Krüger, J. (2019). Deadeye visualization revisited: Investigation of preattentiveness and applicability in virtual environments. *IEEE transactions on visualization and computer graphics*, 26(1):547–557.

[210] Krokos, E., Plaisant, C., and Varshney, A. (2019). Virtual memory palaces: immersion aids recall. *Virtual Reality*, 23(1), doi: 10.1007/s10055-018-0346-3.

[211] Kurillo, G. and Forte, M. (2012). Telearch-integrated visual simulation environment for collaborative virtual archaeology. *Mediterranean Archaeology and Archaeometry*, 12(1):11–20.

[212] Kutsenko, D. (2020). Unity 3D Asset Industrial Scenario Map. `https://assetstore.unity.com/packages/3d/environments/industrial/rpg-fps-/game-assets-for-pc-mobile-industrial-set-v2-0-86679`. Accessed: 2020-08-01.

[213] Kwon, O. H., Muelder, C., Lee, K., and Ma, K. L. (2016). 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.

[214] Laha, B. and Bowman, D. A. (2012). Identifying the benefits of immersion in virtual reality for volume data visualization. In *Immersive visualization revisited workshop of the IEEE VR conference*, pages 1–2.

[215] Laha, B., Sensharma, K., Schiffbauer, J. D., and Bowman, D. A. (2012). Effects of immersion on visual analysis of volume data. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):597–606, doi: 10.1109/TVCG.2012.42.

[216] Lam, H., Bertini, E., Isenberg, P., Plaisant, C., Lam, H., Bertini, E., Isenberg, P., Plaisant, C., and Seven, S. C. (2012). Seven Guiding Scenarios for Information Visualization Evaluation. *Technical Report No. 2011-992-04 Department of Computer Science University of Calgary*.

[217] Langbehn, E., Lubos, P., and Steinicke, F. (2018). Evaluation of locomotion techniques for room-scale VR: Joystick, teleportation, and redirected walking. In *ACM International Conference Proceeding Series*, pages 1–9, doi: 10.1145/3234253.3234291.

[218] Latta, J. N. and Oberg, D. J. (1994). A Conceptual Virtual Reality Model. *IEEE Computer Graphics and Applications*, 14(1):23–29, doi: 10.1109/38.250915.

[219] Lee, B., Bach, B., Dwyer, T., and Marriott, K. (2019). *Immersive Analytics*, volume 39. Springer, 6330 Cham, Switzerland, doi: 10.1109/MCG.2019.2906513.

[220] Lee, B., Brown, D., Lee, B., Hurter, C., Drucker, S., and Dwyer, T. (2020a). Data visceralization: Enabling deeper understanding of data using virtual reality. *arXiv*, doi: 10.1109/tvcg.2020.3030435.

[221] Lee, B., Hu, X., Cordeil, M., Prouzeau, A., Jenny, B., and Dwyer, T. (2020b). Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *arXiv*, doi: 10.1109/tvcg.2020.3030450.

[222] Lee, L., Romano, R., and Stein, G. (2000). Monitoring activities from multiple video streams: Establishing a common coordinate frame. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):758–767, doi: 10.1109/34.868678.

[223] Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, 10(3):282–297, doi: 10.1162/105474601300343612.

[224] Li, X. and Loy, C. C. (2018). Video Object Segmentation with Joint Re-identification and Attention-Aware Mask Propagation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11207 LNCS of *Lecture Notes in Computer Science*, pages 93–110. Springer, doi: 10.1007/978-3-030-01219-9_6.

[225] Lim, S. N., Elgammal, A., and Davis, L. S. (2003). Image-based pan-tilt camera control in a multi-camera surveillance environment. In *Proceedings - IEEE International Conference on Multimedia and Expo*, volume 1, pages 645–648. IEEE, doi: 10.1109/ICME.2003.1221000.

[226] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8693 LNCS(PART 5):740–755, doi: 10.1007/978-3-319-10602-1_48.

[227] Liu, J., Prouzeau, A., Ens, B., and Dwyer, T. (2020). Design and Evaluation of Interactive Small Multiples Data Visualisation in Immersive Spaces. In *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*, pages 588–597. IEEE, doi: 10.1109/VR46266.2020.1581122519414.

[228] Logitech (2020). Logitech BRIDGE. `https://github.com/Logitech/logi_bridge_sdk`. Accessed: 25-May-2019.

[229] Loomis, J. M., Blascovich, J. J., and Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers*, 31(4):557–564, doi: 10.3758/BF03200735.

[230] Luboschik, M., Berger, P., and Staadt, O. (2016). On spatial perception issues in augmented reality based immersive analytics. In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, ISS Companion '16, pages 47–53, New York, NY, USA. ACM, doi: 10.1145/3009939.3009947.

[231] Ma, J., Rosas, L. G., Snowden, M., Wandell, B., Bailenson, J., Greenleaf, W., Dagum, P., Lavori, P., Suppes, T., Lewis, M., Smyth, J., and Williams, L. (2017). Mechanistic Self-Regulation Targets in Integrated Behavior Therapy for Obese and Depressed Adults: Rainbow-Engage Study. In *Annals of Behavioral Medicine*, volume 51, pages S1596–S1597. Springer.

[232] MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., and Hetzler, E. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science*, 32(3):139–160, doi: 10.1559/1523040054738936.

[233] Macik, M. (2018). Cognitive aspects of spatial orientation. *Acta Polytechnica Hungarica*, 15(5):149–167, doi: 10.12700/APH.15.5.2018.5.9.

[234] Mackinlay, J. (1986). Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics (TOG)*, 5(2):110–141, doi: 10.1145/22949.22950.

[235] Madole, D. and Begault, D. (1995). *3-D Sound for Virtual Reality and Multimedia*, volume 19. Academic Press Professional, Inc., San Diego, CA, USA, doi: 10.2307/3680997.

[236] Maes, A., Martinez, X., Druart, K., Laurent, B., Guégan, S., Marchand, C. H., Lemaire, S. D., and Baaden, M. (2018). Minomics, an integrative and immersive tool for multi-omics analysis. *Journal of integrative bioinformatics*.

[237] Maguire, E. A., Burgess, N., and O'Keefe, J. (1999). Human spatial navigation: Cognitive maps, sexual dimorphism, and neural substrates. *Current Opinion in Neurobiology*, 9(2):171–177, doi: 10.1016/S0959-4388(99)80023-3.

[238] Malik, A., Maciejewski, R., Towers, S., McCullough, S., and Ebert, D. S. (2014). Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1863–1872, doi: 10.1109/TVCG.2014.2346926.

[239] McCormack, J., Roberts, J. C., Bach, B., Freitas, C. D. S., Itoh, T., Hurter, C., and Marriott, K. (2018). Multisensory immersive analytics. In Marriott, K. and Dwyer, T. (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11190 LNCS, pages 57–94. Springer, doi: 10.1007/978-3-030-01388-2_3.

[240] McGill, M., Brewster, S., McGookin, D., and Wilson, G. (2020). Acoustic Transparency and the Changing Soundscape of Auditory Mixed Reality. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–16, doi: 10.1145/3313831.3376702.

[241] McGloin, R., Farrar, K., and Krcmar, M. (2013). Video Games, Immersion, and Cognitive Aggression: Does the Controller Matter? *Media Psychology*, 16(1):65–87, doi: 10.1080/15213269.2012.752428.

[242] McIntire, J. P. and Liggett, K. K. (2015). The (possible) utility of stereoscopic 3D displays for information visualization: The good, the bad, and the ugly. In *2014 IEEE VIS International Workshop on 3DVis, 3DVis 2014*, pages 1–9, doi: 10.1109/3DVis.2014.7160093.

[243] Mennecke, B. E., Triplett, J. L., Hassall, L. M., and Conde, Z. J. (2010). Embodied social presence theory. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, pages 1–10. IEEE, doi: 10.1109/HICSS.2010.179.

[244] Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., and Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers and Education*, 70:29–40, doi: 10.1016/j.compedu.2013.07.033.

[245] Merino, L., Fuchs, J., Blumenschein, M., Anslow, C., Ghafari, M., Nierstrasz, O., Behrisch, M., and Keim, D. A. (2017). On the Impact of the Medium in the Effectiveness of 3D Software Visualizations. In *Proceedings - 2017 IEEE Working Conference on Software Visualization, VISSOFT 2017*, volume 2017-October, pages 11–21. IEEE, doi: 10.1109/VISSOFT.2017.17.

[246] Merino, L., Schwarzl, M., Kraus, M., Sedlmair, M., Schmalstieg, D., and Weiskopf, D. (2020). Evaluating mixed and augmented reality: A systematic literature review (2009-2019). In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 438–451. IEEE, doi: 10.1109/ISMAR50242.2020.00069.

[247] Mestre, D. and Fuchs, P. (2006). Immersion et présence. *Le traité de la réalité virtuelle*, pages 309–38.

[248] Milgram, P. and Kishino, F. (1994). Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12):1321–1329.

[249] Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. International Society for Optics and Photonics.

[250] Miller, M., Häußler, J., Kraus, M., Keim, D. A., and El-Assady, M. (2018). Analyzing visual mappings of traditional and alternative music notation. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS 2018*. Retrieved from http://arxiv.org/abs/1810.10814.

[251] Miller, M., Schäfer, H., Kraus, M., Leman, M., Keim, D. A., and El-Assady, M. (2019). Framing visual musicology through methodology transfer. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS 2019*. Retrieved from http://arxiv.org/abs/1908.10411.

[252] Mills, J. W. (2008). *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques, and Software Tools, Second Edition - by Michael J. de Smith, Michael F. Goodchild, and Paul A. Longley*, volume 12. Troubador Publishing Ltd, doi: 10.1111/j.1467-9671.2008.01122.x.

[253] Moffat, S. D., Zonderman, A. B., and Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, 22(5):787–796, doi: 10.1016/S0197-4580(01)00251-2.

[254] Morgan, B. J. T. and Ray, A. P. G. (1995). Non-uniqueness and Inversions in Cluster Analysis. *Applied Statistics*, 44(1):117, doi: 10.2307/2986199.

[255] Mota, R. C., Rocha, A., Silva, J. D., Alim, U., and Sharlin, E. (2018). 3de interactive lenses for visualization in virtual environments. In *2018 IEEE Scientific Visualization Conference (SciVis)*, pages 21–25. IEEE.

[256] Mount, N., Chambers, C., Weaver, D., and Priestnall, G. (2009). Learner immersion engagement in the 3D virtual world: principles emerging from the DELVE project. *Innovation in Teaching and Learning in Information and Computer Sciences*, 8(3):40–55, doi: 10.11120/ital.2009.08030040.

[257] Müller, J., Butscher, S., Feyer, S. P., and Reiterer, H. (2017). Studying collaborative object positioning in distributed augmented realities. In *ACM International Conference Proceeding Series*, pages 123–132. ACM Press, doi: 10.1145/3152832.3152856.

[258] Munzner, T. (2008). Process and pitfalls in writing information visualization research papers. In Kerren, A., Stasko, J. T., Fekete, J.-D., and North, C. (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4950 LNCS, pages 134–153. Springer-Verlag, Berlin, Heidelberg, doi: 10.1007/978-3-540-70956-5_6.

[259] Munzner, T. (2015). *Visualization Analysis & Design*. CRC Press.

[260] Nagel, H. R., Granum, E., Bovbjerg, S., and Vittrup, M. (2008). *Immersive Visual Data Mining: The 3DVDM Approach*, pages 281–311. Springer Berlin Heidelberg, Berlin, Heidelberg, doi: 10.1007/978-3-540-71080-6_18.

[261] Nagel, H. R., Granum, E., and Musaeus, P. (2001). Methods for visual mining of data in virtual reality. In *Proceedings of the International Workshop on Visual Data Mining*, pages 13–27.

[262] Nakaya, T. and Yano, K. (2010). Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. *Transactions in GIS*, 14(3):223–239, doi: 10.1111/j.1467-9671.2010.01194.x.

[263] Naudts, L. (2020). Criminal Profiling and Non-Discrimination: On Firm Grounds for the Digital Era? *Security and Law*, pages 63–96, doi: 10.1017/9781780688909.004.

[264] Naudts, L. and Vogiatzoglou, P. (2019). The VICTORIA Ethical and Legal Management Toolkit: Practice and Theory for Video Surveillance.

[265] Nebeling, M., Speicher, M., Wang, X., Rajaram, S., Hall, B. D., Xie, Z., Raistrick, A. R., Aebersold, M., Happ, E. G., Wang, J., et al. (2020). Mrat: The mixed reality analytics toolkit. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

[266] Nelson, L., Cook, D., and Cruz-Neira, C. (1999). XGobi vs the C2: Results of an experiment comparing data visualization in a 3-D immersive virtual reality environment with a 2-D workstation display. *Computational Statistics*, 14(1):39–51, doi: 10.1007/pl00022704.

[267] Noack, R. (2016). Leaked document says 2,000 men allegedly assaulted 1,200 German women on New Year's Eve. *The Washington Post*.

[268] Noel, S., Jacobs, M., Kalapa, P., and Jajodia, S. (2005). Multiple coordinated views for network attack graphs. In *IEEE Workshop on Visualization for Computer Security 2005, VizSEC 05, Proceedings*, pages 99–106, Minneapolis, MN, USA. IEEE, doi: 10.1109/VIZSEC.2005.1532071.

[269] Norrby, M., Grebner, C., Eriksson, J., and Boström, J. (2015). Molecular Rift: Virtual Reality for Drug Designers. *Journal of Chemical Information and Modeling*, 55(11):2475–2484, doi: 10.1021/acs.jcim.5b00544.

[270] Nowke, C., Schmidt, M., Van Albada, S. J., Eppler, J. M., Bakker, R., Diesrnann, M., Hentschel, B., and Kuhlen, T. (2013). VisNEST - Interactive analysis of neural activity data. In *BioVis 2013 - IEEE Symposium on Biological Data Visualization 2013, Proceedings*, pages 65–72, doi: 10.1109/BioVis.2013.6664348.

[271] O'Connor, M. B., Bennie, S. J., Deeks, H. M., Jamieson-Binnie, A., Jones, A. J., Shannon, R. J., Walters, R., Mitchell, T. J., Mulholland, A. J., and Glowacki, D. R. (2019). Interactive molecular dynamics in virtual reality from quantum chemistry to drug binding: An open-source multi-person framework. *Journal of Chemical Physics*, 150(22):220901, doi: 10.1063/1.5092590.

[272] Oneto, L. and Chiappa, S. (2020). Fairness in Machine Learning. In *Studies in Computational Intelligence*, volume 896, pages 155–196. Springer, doi: 10.1007/978-3-030-43883-8_7.

[273] OpenSpace (2019). OpenSpace Computer Program. `https://www.openspaceproject.com/`. Accessed: 2019-09-17.

[274] Pahins, C. A., Stephens, S. A., Scheidegger, C., and Comba, J. L. (2017). Hashedcubes: Simple, Low Memory, Real-Time Visual Exploration of Big Data. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):671–680, doi: 10.1109/TVCG.2016.2598624.

[275] Palomo, C., Guo, Z., Silva, C. T., and Freire, J. (2016). Visually Exploring Transportation Schedules. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):170–179, doi: 10.1109/TVCG.2015.2467592.

[276] Papadopoulos, C., Gutenko, I., and Kaufman, A. E. (2016). VEEVVIE: Visual Explorer for Empirical Visualization, VR and Interaction Experiments. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):111–120, doi: 10.1109/TVCG.2015.2467954.

[277] Park, H. S., Yoon, J. W., Kim, J., Iseki, K., and Hallett, M. (2011). Development of a VR-based treadmill control interface for gait assessment of patients with Parkinson's disease. In *IEEE International Conference on Rehabilitation Robotics*, pages 1–5. IEEE, doi: 10.1109/ICORR.2011.5975463.

[278] Patnaik, B., Batch, A., and Elmqvist, N. (2019). Information olfactation: Harnessing scent to convey data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):726–736, doi: 10.1109/TVCG.2018.2865237.

[279] Patrão, B., Pedro, S., and Menezes, P. (2020). How to Deal with Motion Sickness in Virtual Reality. In Dias, P. and Menezes, P. (Eds.), *Portuguese Meeting on Computer Graphics and Interaction 2015*. The Eurographics Association, doi: 10.2312/pt.20151201.

[280] Pausch, R., Crea, T., and Conway, M. (1992). A Literature Survey for Virtual Environments: Military Flight Simulator Visual Systems and Simulator Sickness. *Presence: Teleoperators and Virtual Environments*, 1(3):344–363, doi: 10.1162/pres.1992.1.3.344.

[281] Perin, C., Vuillemot, R., and Fekete, J. D. (2013). SoccerStories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, doi: 10.1109/TVCG.2013.192.

[282] Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., and Richardson, C. (1995). A redrawn vandenberg and kuse mental rotations test - different versions and factors that affect performance. *Brain and Cognition*, 28(1):39–58, doi: 10.1006/brcg.1995.1032.

[283] Peterson, S., Axholt, M., and Ellis, S. R. (2008). Managing visual clutter: A generalized technique for label segregation using stereoscopic disparity. In *Proceedings - IEEE Virtual Reality*, pages 169–176. IEEE, doi: 10.1109/VR.2008.4480769.

[284] Pham, T. T. T., Le, T. L., Vu, H., Dao, T. K., and Nguyen, V. T. (2017). Fully-automated person re-identification in multi-camera surveillance system with a robust kernel descriptor and effective shadow removal method. *Image and Vision Computing*, 59:44–62, doi: 10.1016/j.imavis.2016.10.010.

[285] Pilgrim, M., Bouchlaghem, D., Loveday, D., and Holmes, M. (2000). Abstract data visualisation in the built environment. In *2000 IEEE Conference on Information Visualization. An International Conference on Computer Visualization and Graphics*, pages 126–134.

[286] Piringer, H., Kosara, R., and Hauser, H. (2004). Interactive focus+context visualization with linked 2D/3D scatterplots. In *Proceedings - Second International Conference on Coordinated and Multiple Views in Exploratory Visualization, CMV 2004*, pages 49–60, doi: 10.1109/CMV.2004.1319526.

[287] Poco, J., Mayhua, A., and Heer, J. (2018). Extracting and Retargeting Color Mappings from Bitmap Images of Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):637–646, doi: 10.1109/TVCG.2017.2744320.

[288] Pollok, T. (2018). A New Multi-Camera Dataset with Surveillance, Mobile and Stereo Cameras for Tracking, Situation Analysis and Crime Scene Investigation Applications. In *Proceedings of the 2018 the 2Nd International Conference on Video and Image Processing*, ICVIP 2018, pages 171–175, doi: 10.1145/3301506.3301542.

[289] Pollok, T., Kraus, M., Qu, C., Miller, M., Moritz, T., Kilian, T., Keim, D., and Jentner, W. (2019). Computer vision meets visual analytics: Enabling 4D crime scene investigation from image and video data. In *IET Conference Publications*, volume 2019, pages 44–49, doi: 10.1049/cp.2019.1166.

[290] Polys, N. F. and Bowman, D. A. (2004). Design and display of enhancing information in desktop information-rich virtual environments: Challenges and techniques. *Virtual Reality*, 8(1):41–54, doi: 10.1007/s10055-004-0134-0.

[291] Porter, J., Anand, T., Johnson, B., Khan, R. M., and Sobel, N. (2005). Brain mechanisms for extracting spatial information from smell. *Neuron*, 47(4):581–592, doi: 10.1016/j.neuron.2005.06.028.

[292] Portman, M. E., Natapov, A., and Fisher-Gewirtzman, D. (2015). To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. *Computers, Environment and Urban Systems*, 54:376–384, doi: 10.1016/j.compenvurbsys.2015.05.001.

[293] Prabhat, Forsberg, A., Katzourin, M., Wharton, K., and Slater, M. (2008). A comparative study of Desktop, fishtank, and cave systems for the exploration of volume rendered confocal data sets. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):551–563, doi: 10.1109/TVCG.2007.70433.

[294] Probst, D. and Reymond, J. L. (2018). Exploring DrugBank in Virtual Reality Chemical Space. *Journal of Chemical Information and Modeling*, 58(9):1731–1735, doi: 10.1021/acs.jcim.8b00402.

[295] Prouzeau, A., Cordeil, M., Robin, C., Ens, B., Thomas, B. H., and Dwyer, T. (2019). Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3D scatterplots. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–12, doi: 10.1145/3290605.3300555.

[296] Pryke, A., Mostaghim, S., and Nazemi, A. (2007). Heatmap visualization of population based multi objective algorithms. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4403 LNCS, pages 361–375, Berlin, Heidelberg. Springer, doi: 10.1007/978-3-540-70928-2_29.

[297] Qin, H., Patrick Rau, P. L., and Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human-Computer Interaction*, 25(2):107–133, doi: 10.1080/10447310802546732.

[298] Qu, C., Metzler, J., and Monari, E. (2018). IvisX: An integrated video investigation suite for forensic applications. In *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2018*, volume 2018-January, pages 9–17. IEEE Computer Society, doi: 10.1109/WACVW.2018.00007.

[299] Rademacher, T. (2019). Artificial intelligence and law enforcement. In *Regulating Artificial Intelligence*, pages 225–254. Springer, doi: 10.1007/978-3-030-32361-5_10.

[300] Ragan, E. D., Sowndararajan, A., Kopper, R., and Bowman, D. (2010). The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments. *Presence: Teleoperators and Virtual Environments*, 19(6):527–543, doi: 10.1162/pres_a_00016.

[301] Raja, D., Bowman, D. A., Lucas, J., and North, C. (2004). Exploring the Benefits of Immersion in Abstract Information Visualization. In *Proceedings of the 8th Immersive Projection Technology Workshop*, pages 61–69, doi: 10.1.1.205.235.

[302] Rau, R., Bohk-Ewald, C., Muszyńska, M. M., and Vaupel, J. W. (2018). *Introduction: Why Do We Visualize Data and What Is This Book About?*, pages 1–4. Springer International Publishing, Cham, doi: 10.1007/978-3-319-64820-0_1.

[303] Razzaque, S., Kohn, Z., and Whitton, M. C. (2001). Redirected Walking. In *Proceedings of EUROGRAPHICS*, pages 289–294. Eurographics Association, doi: 10.2312/egs.20011036.

[304] Reda, K., Febretti, A., Knoll, A., Aurisano, J., Leigh, J., Johnson, A., Papka, M. E., and Hereld, M. (2013). Visualizing large, heterogeneous data in hybrid-reality environments. *IEEE Computer Graphics and Applications*, 33(4):38–48.

[305] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2016-December, pages 779–788, doi: 10.1109/CVPR.2016.91.

[306] Redmon, J. and Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, volume 2017-January, pages 6517–6525, doi: 10.1109/CVPR.2017.690.

[307] Redmon, J. and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *CoRR*.

[308] Regenbrecht, H. and Schubert, T. (2002). Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 11(4):425–434, doi: 10.1162/105474602760204318.

[309] Rehm, F., Klawonn, F., and Kruse, R. (2007). A novel approach to noise clustering for outlier detection. *Soft Computing*, 11(5):489–494, doi: 10.1007/s00500-006-0112-4.

[310] Reina, G. and Ertl, T. (2004). Volume Visualization and Visual Queries for Large High-Dimensional Datasets. In *Proceedings of the Joint Eurographics – IEEE TCVG Symposium on Visualization*, pages 255–260, doi: 10.2312/VisSym/VisSym04/255-260.

[311] Reipschläger, P., Flemisch, T., and Dachselt, R. (2020). Personal augmented reality for information visualization on large interactive displays. *arXiv*, 27(1):1, doi: 10.1109/tvcg.2020.3030460.

[312] Rhyne, T. M., MacEachren, A. M., Gahegan, M., Pike, W., Brewer, I., Cai, G., Lengerich, E., and Hardisty, F. (2004). Geovisualization for Knowledge Construction and Decision Support. *IEEE Computer Graphics and Applications*, 24(1):13–17, doi: 10.1109/MCG.2004.1255801.

[313] Riecke, B. E. and Schulte-Pelkum, J. (2013). Perceptual and cognitive factors for self-motion simulation in virtual environments: How can self-motion illusions (Vection) be utilized? In *Human Walking in Virtual Environments: Perception, Technology, and Applications*, volume 9781441984326, pages 27–54. Springer, doi: 10.1007/978-1-4419-8432-6_2.

[314] Riva, G., Waterworth, J. A., Waterworth, E. L., and Mantovani, F. (2011). From intention to action: The role of presence. *New Ideas in Psychology*, 29(1):24–37, doi: 10.1016/j.newideapsych.2009.11.002.

[315] roadtovr.com (2021). Road to VR - Hardware Review Archives. `https://www.roadtovr.com/category/hardware-review/`. Accessed: 2021-10-01.

[316] Rodriguez, H., Beck, D., Lind, D., and Lok, B. (2008). Audio analysis of human/virtual-human interaction. In Prendinger, H., Lester, J., and Ishizuka, M. (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5208 LNAI, pages 154–161, Berlin, Heidelberg. Springer Berlin Heidelberg, doi: 10.1007/978-3-540-85483-8_16.

[317] Rossion, B. and Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2):217–236, doi: 10.1068/p5117.

[318] Ruddle, R. A. (2005). The effect of trails on first-time and subsequent navigation in a virtual environment. In *Proceedings - IEEE Virtual Reality*, pages 115–122. IEEE, IEEE, doi: 10.1109/vr.2005.76.

[319] Ruddle, R. A., Payne, S. J., and Jones, D. M. (1999). Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence: Teleoperators and Virtual Environments*, 8(2):157–168, doi: 10.1162/105474699566143.

[320] Sacha, D., Jentner, W., Zhang, L., Stoffel, F., Ellis, G., and Keim, D. (2017). Applying Visual Interactive Dimensionality Reduction to Criminal Intelligence Analysis. *VALCRI White Paper Series*, 1.

[321] Sacha, D., Kraus, M., Keim, D. A., and Chen, M. (2018). Vis4ml: An ontology for visual analytics assisted machine learning. *IEEE transactions on visualization and computer graphics*, 25(1):385–395, doi: 10.1109/TVCG.2018.2864838.

[322] Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., and Keim, D. A. (2014). Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604–1613, doi: 10.1109/TVCG.2014.2346481.

[323] Saenz, M., Baigelenov, A., Hung, Y.-H., and Parsons, P. (2017). Reexamining the cognitive utility of 3D visualizations using augmented reality holograms. *Workshop on Immersive Analytics: Exploring Future Interaction and Visualization Technologies for Data Analytics (Immersive 2017)*, (October).

[324] Sancar, F. H. (1986). *Wayfinding in Architecture*, volume 5 of *Environmental Design Series (Stroudsburg, Pa.), V. 4*. Van Nostrand Reinhold, doi: 10.3368/lj.5.1.71.

[325] Sarafianos, N., Boteanu, B., Ionescu, B., and Kakadiaris, I. A. (2016). 3D Human pose estimation: A review of the literature and analysis of covariates. *Computer Vision and Image Understanding*, 152:1–20, doi: 10.1016/j.cviu.2016.09.002.

[326] Sarkar, M. and Brown, M. H. (1992). Graphical fisheye views of graphs. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 83–92, doi: 10.1145/142750.142763.

[327] Sathiyanarayanan, M., Junejo, A. K., and Fadahunsi, O. (2019). Visual Analysis of Predictive Policing to Improve Crime Investigation. In *Proceedings of the 4th International Conference on Contemporary Computing and Informatics, IC3I 2019*, pages 197–203. IEEE, doi: 10.1109/IC3I46837.2019.9055515.

[328] Sawant, N., Scharver, C., Leigh, J., Johnson, A., Reinhart, G., Creel, E., Batchu, S., Bailey, S., and Grossman, R. (2000). The tele-immersive data explorer: A distributed architecture for collaborative interactive visualization of large data-sets. In *4th International Immersive Projection Technology Workshop, Ames, Iowa*.

[329] Schmidt, S. (2012). Max Planck Institute for Chemistry. `https://www.movebank.org/`. Accessed: 2019-05-25.

[330] Schonberger, J. L. and Frahm, J. M. (2016). Structure-from-Motion Revisited. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2016-December, pages 4104–4113. IEEE Computer Society, doi: 10.1109/CVPR.2016.445.

[331] Schreck, T., Bernard, J., Tekušová, T., and Kohlhammer, J. (2008). Visual cluster analysis of trajectory data with interactive Kohonen maps. *VAST'08 - IEEE Symposium on Visual Analytics Science and Technology, Proceedings*, 8(1):3–10, doi: 10.1109/VAST.2008.4677350.

[332] Schroeder, K., Ajdadilish, B., Henkel, A. P., and Calero Valdez, A. (2020). Evaluation of a Financial Portfolio Visualization using Computer Displays and Mixed Reality Devices with Domain Experts. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–9, doi: 10.1145/3313831.3376556.

[333] Schubring, D., Kraus, M., Stolz, C., Weiler, N., Keim, D. A., and Schupp, H. T. (2020). Virtual reality potentiates emotion and task effects of alpha/beta brain oscillations. *Brain Sciences*, 10(8):537, doi: 10.3390/brainsci10080537.

[334] Schuchardt, P. and Bowman, D. A. (2007). The benefits of immersion for spatial understanding of complex underground cave systems. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, pages 121–124. ACM, doi: 10.1145/1315184.1315205.

[335] Sedlmair, M., Munzner, T., and Tory, M. (2013). Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643, doi: 10.1109/TVCG.2013.153.

[336] Sedlmair, M., Tatu, A., Munzner, T., and Tory, M. (2012). A taxonomy of visual cluster separation factors. *Computer Graphics Forum*, 31(3 PART 4):1335–1344, doi: 10.1111/j.1467-8659.2012.03125.x.

[337] Seibert, J. and Shafer, D. M. (2018). Control mapping in virtual reality: effects on spatial presence and controller naturalness. *Virtual Reality*, 22(1):79–88, doi: 10.1007/s10055-017-0316-1.

[338] Shneiderman, B. (1987). *Designing the user interface strategies for effective human-computer interaction*, volume 9. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 3rd edition, doi: 10.1145/25065.950626.

[339] Shneiderman, B. (1996). Eyes have it: a task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages, Proceedings*, pages 336–343, Boulder, CO, USA. IEEE, doi: 10.1109/vl.1996.545307.

[340] Shneiderman, B. and Maes, P. (1997). Direct manipulation vs. interface agents. *Interactions*, 4(6):42–61, doi: 10.1145/267505.267514.

[341] Sicat, R., Li, J., Choi, J., Cordeil, M., Jeong, W. K., Bach, B., and Pfister, H. (2019). DXR: A Toolkit for Building Immersive Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):715–725, doi: 10.1109/TVCG.2018.2865152.

[342] Simeone, A. L., Velloso, E., and Gellersen, H. (2015). Substitutional reality: Using the physical environment todesign virtual reality experiences. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 2015-April, pages 3307–3316. ACM, doi: 10.1145/2702123.2702389.

[343] Skarbez, R., Polys, N. F., Ogle, J. T., North, C., and Bowman, D. A. (2019). Immersive analytics: Theory and research agenda. *Frontiers in Robotics and AI*, 6:82.

[344] Skeels, M., Lee, B., Smith, G., and Robertson, G. G. (2010). Revealing uncertainty for information visualization. *Information Visualization*, 9(1):70–81, doi: 10.1057/ivs.2009.1.

[345] Slater, M. (2003). A note on presence. *Presence Connect*, 3(January 2003):1–5.

[346] Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364:3549–3557.

[347] Slater, M., Lotto, B., Arnold, M. M., and Sanchez-Vives, M. V. (2009). How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia*, 40(2):193–210, doi: 1988-5253.

[348] Slater, M. and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6):603–616, doi: 10.1162/pres.1997.6.6.603.

[349] Slay, H., Phillips, M., Vernik, R., and Thomas, B. (2001). Interaction modes for augmented reality visualization. In *Proceedings of the 2001 Asia-Pacific Symposium on Information Visualisation - Volume 9*, APVis '01, pages 71–75, Darlinghurst, Australia, Australia. Australian Computer Society, Inc.

[350] Sloot, L. H., van der Krogt, M. M., and Harlaar, J. (2014). Effects of adding a virtual reality environment to different modes of treadmill walking. *Gait and Posture*, 39(3):939–945, doi: 10.1016/j.gaitpost.2013.12.005.

[351] Smith, N. G., Knabb, K., Defanti, C., Weber, P., Schulze, J., Prudhomme, A., Kuester, F., Levy, T. E., and Defanti, T. A. (2013). ArtifactVis2: Managing real-time archaeological data in immersive 3D environments. In *Proceedings of the DigitalHeritage 2013 - Federating the 19th Int'l VSMM, 10th Eurographics GCH, and 2nd UNESCO Memory of the World Conferences, Plus Special Sessions fromCAA, Arqueologica 2.0 et al.*, volume 1, pages 363–370, doi: 10.1109/DigitalHeritage.2013.6743761.

[352] Sommer, B., Barnes, D. G., Boyd, S., Chandler, T., Cordeil, M., Czauderna, T., Klapperstück, M., Klein, K., Nguyen, T. D., Nim, H., et al. (2017). 3d-stereoscopic immersive analytics projects at monash university and university of konstanz. *Electronic Imaging*, 2017(5):179–187.

[353] Sommer, B., Bender, C., Hoppe, T., Gamroth, C., and Jelonek, L. (2014). Stereoscopic cell visualization: from mesoscopic to molecular scale. In *Electronic Imaging, Proceedings of Stereoscopic Displays and Applications XXVIII*, volume 23, pages 011007–1 – 011007–10, doi: 10.1117/1.JEI.23.1.011007.

[354] Sommer, B., Wang, S. J., Xu, L., Chen, M., and Schreiber, F. (2015). Hybrid-dimensional visualization and interaction-integrating 2d and 3d visualization with semi-immersive navigation techniques. In *2015 Big Data Visual Analytics (BDVA)*, pages 1–8. IEEE.

[355] Steck, S. D. and Mallot, H. A. (2000). The role of global and local landmarks in virtual environment navigation. *Presence: Teleoperators and Virtual Environments*, 9(1):69–83, doi: 10.1162/105474600566628.

[356] Stein, M., Janetzko, H., Breitkreutz, T., Seebacher, D., Schreck, T., Grossniklaus, M., Couzin, I. D., and Keim, D. A. (2016). Director's Cut: Analysis and Annotation of Soccer Matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, doi: 10.1109/MCG.2016.102.

[357] Stein, M., Janetzko, H., Lamprecht, A., Breitkreutz, T., Zimmermann, P., Goldlücke, B., Schreck, T., Andrienko, G., Grossniklaus, M., and Keim, D. A. (2018). Bring It to the Pitch: Combining Video and Movement Data to Enhance Team Sport Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, doi: 10.1109/TVCG.2017.2745181.

[358] Stellmach, S., Nacke, L., and Dachselt, R. (2010). Advanced gaze visualizations for three-dimensional virtual environments. In *Eye Tracking Research and Applications Symposium (ETRA)*, pages 109–112. ACM, doi: 10.1145/1743666.1743693.

[359] Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42(4):73–93, doi: 10.1111/j.1460-2466.1992.tb00812.x.

[360] Stolk, B., Abdoelrahman, F., Koning, A., Wielinga, P., Neefs, J.-M., Stubbs, A., de Bondt, A., Leemans, P., and van der Spek, P. (2002). Mining the human genome using virtual reality. In *Proceedings of the Fourth Eurographics Workshop on Parallel Graphics and Visualization*, EGPGV '02, pages 17–21, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

[361] Strandburg-Peshkin, A., Farine, D. R., Crofoot, M. C., and Couzin, I. D. (2017). Habitat and social factors shape individual decisions and emergent group structure during baboon collective movement. *eLife*, 6:e19505, doi: 10.7554/eLife.19505.

[362] Stuerzlinger, W. and Wingrave, C. A. (2011). The value of constraints for 3d user interfaces. In *Virtual Realities: Dagstuhl Seminar 2008*, pages 203–223, doi: 10.1007/978-3-211-99178-7_11.

[363] Suzuki, K., Wakisaka, S., and Fujii, N. (2012). Substitutional reality system: A novel experimental platform for experiencing alternative reality. *Scientific Reports*, 2:1–9, doi: 10.1038/srep00459.

[364] Symanzik, J., Cook, D., Kohlmeyer, B. D., Lechner, U., and Cruz-neira, C. (1997). Dynamic statistical graphics in the c2 virtual reality environment. *Computing Science and Statistics*, 29:41–47.

[365] Teoh, S. T., Ma, K. L., Wu, S. F., and Zhao, X. (2002). Case study: Interactive visualization for internet security. In *Proceedings of the IEEE Visualization Conference*, pages 505–508, doi: 10.1109/visual.2002.1183816.

[366] Thomas, J. J. and Cook, K. A. (2006). A visual analytics agenda. *IEEE Comput. Graph. Appl.*, 26(1):10–13, doi: 10.1109/MCG.2006.5.

[367] Thompson, P. M., Hayashi, K. M., De Zubicaray, G., Janke, A. L., Rose, S. E., Semple, J., Herman, D., Hong, M. S., Dittmer, S. S., Doddrell, D. M., and Toga, A. W. (2003). Dynamics of gray matter loss in Alzheimer's disease. *Journal of Neuroscience*, 23(3):994–1005, doi: 10.1523/jneurosci.23-03-00994.2003.

[368] Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, 8(3):115–121, doi: 10.1016/j.tics.2004.01.006.

[369] Tory, M., Sprague, D. W., Wu, F., So, W. Y., and Munzner, T. (2007). Spatialization design: Comparing points and landscapes. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1262–1269, doi: 10.1109/TVCG.2007.70596.

[370] Tunguz, T. (2018). The Cost Of Bad Data Is The Illusion Of Knowledge. `http://tomtunguz.com/cost-of-bad-data-1-10-100/`. Accessed: 2019-05-25.

[371] Unity (2020). Unity 3D Developer Website. `https://www.unity3d.com/`. Accessed: 2020-08-17.

[372] USTWO (2018). Blog Entry - USTWO Designing for VR Google Carboard. `https://www.ustwo.com/blog/designing-for-virtual-reality-google-cardboard/`. Accessed: 2018-07-19.

[373] Valve (2020). Valve Index Developer Website. `https://store.steampowered.com/valveindex/`. Accessed: 2020-07-30.

[374] Van Dam, A., Forsberg, A. S., Laidlaw, D. H., LaViola, J. J., and Simpson, R. M. (2000). Immersive VR for scientific visualization: A progress report. *IEEE Computer Graphics and Applications*, 20(6):26–52.

[375] Van Der Meijden, O. A. and Schijven, M. P. (2009). The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: A current review. *Surgical Endoscopy*, 23(6):1180–1190, doi: 10.1007/s00464-008-0298-x.

[376] Van Schooten, B. W., Van Dijk, E. M., Zudilova-Seinstra, E., Suinesiaputra, A., and Reiber, J. H. (2010). The effect of stereoscopy and motion cues on 3D interpretation task performance. In *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, pages 167–170. ACM, doi: 10.1145/1842993.1843023.

[377] van Teylingen, R., Ribarsky, W., and van der Mast, C. (1997). Virtual data visualizer. *IEEE Transactions on Visualization and Computer Graphics*, 3(1):65–74, doi: 10.1109/2945.582350.

[378] Van Veen, H. A., Distler, H. K., Braun, S. J., and Bülthoff, H. H. (1998). Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Computer Systems*, 14(3-4):231–242, doi: 10.1016/s0167-739x(98)00027-2.

[379] Van Walsum, T., Post, F. H., Silver, D., and Post, F. J. (1996). Feature extraction and iconic visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2(2):111–119, doi: 10.1109/2945.506223.

[380] van Wijk, J. (2006). The Value of Visualization. In *IEEE Visualization Conf.*, pages 79–86, doi: 10.1109/visual.2005.1532781.

[381] van Wijk, J. J. and van Selow, E. R. (1999). Cluster and calendar based visualization of time series data. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 4–9. IEEE, doi: 10.1109/infvis.1999.801851.

[382] Verbree, E., Maren, G. V., Germs, R., Jansen, F., and Kraak, M. J. (1999). Interaction in virtual world views-linking 3D GIS with VR. *International Journal of Geographical Information Science*, 13(4):385–396, doi: 10.1080/136588199241265.

[383] VICTORIA (2019). VICTORIA - Video Analysis for Investigation of Criminal and Terrorist Activities. `https://www.victoria-project.eu/`. Accessed: 25-May-2019.

[384] virtualspeech (2021). virtualspeech Blog Entry. `https://virtualspeech.com/blog/vr-applications`. Accessed: 2021-04-08.

[385] VRS (2020). Blog Entry - Virtual Reality Environments. `https://www.vrs.org.uk/virtual-reality-environments/`. Accessed: 2018-07-19.

[386] Wagner Filho, J. A., Freitas, C. M., and Nedel, L. (2018a). VirtualDesk: A Comfortable and Efficient Immersive Information Visualization Approach. *Computer Graphics Forum*, 37(3):415–426, doi: 10.1111/cgf.13430.

[387] Wagner Filho, J. A., Rey, M. F., Freitas, C. M., and Nedel, L. (2018b). Immersive Visualization of Abstract Information: An Evaluation on Dimensionally-Reduced Data Scatterplots. In *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings*, pages 483–490, doi: 10.1109/VR.2018.8447558.

[388] Wagner Filho, J. A., Rey, M. F., Freitas, C. M. D. S., and Nedel, L. (2017). Immersive Analytics of Dimensionally-Reduced Data Scatterplots. *2nd Workshop on Immersive Analytics*.

[389] Wagner Filho, J. A., Stuerzlinger, W., and Nedel, L. (2019). Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):514–524.

[390] Walshe, D. G., Lewis, E. J., Kim, S. I., O'Sullivan, K., and Wiederhold, B. K. (2003). Exploring the use of computer games and virtual reality in exposure therapy for fear of driving following a motor vehicle accident. *Cyberpsychology and Behavior*, 6(3):329–334, doi: 10.1089/109493103322011641.

[391] Wang, H., Lu, Y., Shutters, S. T., Steptoe, M., Wang, F., Landis, S., and MacIejewski, R. (2019a). A visual analytics framework for spatiotemporal trade network analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):331–341, doi: 10.1109/TVCG.2018.2864844.

[392] Wang, T., Hu, X., Wang, Q., Heng, P.-A., and Fu, C.-W. (2020). Instance Shadow Detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1877–1886, doi: 10.1109/cvpr42600.2020.00195.

[393] Wang, X., Chen, W., Chou, J. K., Bryan, C., Guan, H., Chen, W., Pan, R., and Ma, K. L. (2019b). GraphProtector: A Visual Interface for Employing and Assessing Multiple Privacy Preserving Graph Algorithms. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):193–203, doi: 10.1109/TVCG.2018.2865021.

[394] Wang, X. M., Zhang, T. Y., Ma, Y. X., Xia, J., and Chen, W. (2016). A Survey of Visual Analytic Pipelines. *Journal of Computer Science and Technology*, 31(4):787–804, doi: 10.1007/s11390-016-1663-1.

[395] Wanscher, M., Riishede, L., and Krogh, B. (1985). Fistula Formation following Epidural Catheter:A Case Report. In Card, S. K., Mackinlay, J. D., and Shneiderman, B. (Eds.), *Acta Anaesthesiologica Scandinavica*, volume 29, pages 552–553. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, doi: 10.1111/j.1399-6576.1985.tb02252.x.

[396] Ward, M. O., Grinstein, G., and Daniel, K. (2015). *Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition*. CRC Press.

[397] Ware, C. (2004). *Information Visualization: Perception for Design: Second Edition*. Elsevier, San Francisco, 2\textsuperscriptnd edition, doi: 10.1016/B978-1-55860-819-1.X5000-6.

[398] Ware, C. and Franck, G. (1994). Viewing a graph in a virtual reality display is three times as good as a 2d diagram. In *Visual Languages, 1994. Proceedings., IEEE Symposium on*, pages 182–183. IEEE.

[399] Ware, C. and Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics (TOG)*, 15(2):121–140.

[400] Ware, C. and Mitchell, P. (2008). Visualizing graphs in three dimensions. *ACM Transactions on Applied Perception*, 5(1):1–15, doi: 10.1145/1279640.1279642.

[401] Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411):664–675, doi: 10.1080/01621459.1990.10474926.

[402] Weiler, N., Kraus, M., Kilian, T., Jentner, W., and Keim, D. A. (2018). Visual Analytics for Semi-Automatic 4D Crime Scene Reconstruction. In *4th International Symposium on Big Data Visual and Immersive Analytics*.

[403] Whitlock, M., Smart, S., and Szafir, D. A. (2020). Graphical Perception for Immersive Analytics. In *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*, pages 616–625. IEEE, doi: 10.1109/VR46266.2020.1582298687237.

[404] Wickham, H., Hofmann, H., Wickham, C., and Cook, D. (2012). Glyph-maps for visually exploring temporal patterns in climate data and models. *Environmetrics*, 23(5):382–393, doi: 10.1002/env.2152.

[405] Wijayasekara, D., Linda, O., and Manic, M. (2011). Cave-som: Immersive visual data mining using 3d self-organizing maps. In *Neural Networks (IJCNN), The 2011 International Joint Conference on*, pages 2471–2478. IEEE.

[406] Wikelski, M. and Kays, R. (2016). Movebank: archive, analysis and sharing of animal movement data. *Hosted by the Max Planck Institute for Ornithology*.

[407] wikipedia.org (2021). Wikipedia: Comparison of virtual reality headsets. `https://en.wikipedia.org/w/index.php?title=Comparison_of_virtual_reality_headsets&oldid=1047329569`. Accessed: 2021-10-01.

[408] Willems, N., Van De Wetering, H., and Van Wijk, J. J. (2009). Visualization of vessel movements. In *Computer Graphics Forum*, volume 28, pages 959–966, Chichester, UK. Wiley Online Library, IEEE, doi: 10.1111/j.1467-8659.2009.01440.x.

[409] Willett, W., Jansen, Y., and Dragicevic, P. (2016). Embedded data representations. *IEEE transactions on visualization and computer graphics*, 23(1):461–470.

[410] Williams, B., Narasimham, G., Rump, B., McNamara, T. P., Carr, T. H., Rieser, J., and Bodenheimer, B. (2007). Exploring large virtual environments with an HMD when physical space is limited. In *ACM International Conference Proceeding Series*, volume 253, pages 41–48. ACM, doi: 10.1145/1272582.1272590.

[411] Wills, S. and Roecker, S. (2017). Statistics for Soil Survey. http://ncss-tech.github.io/stats_for_soil_survey/chapters/9_uncertainty/Uncert_val.html#1_introduction. Accessed: 2019-05-25.

[412] Wither, J. and Höllerer, T. (2005). Pictorial depth cues for outdoor augmented reality. In *Proceedings - International Symposium on Wearable Computers, ISWC*, volume 2005, pages 92–99, doi: 10.1109/ISWC.2005.41.

[413] Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, doi: 10.1162/105474698565686.

[414] Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of International Conference on Evaluation and Assessment in Software Engineering*, pages 1–10. ACM.

[415] Wong, P. C. and Thomas, J. (2004). Visual analytics. *IEEE Computer Graphics and Applications*, 24(5):20–21, doi: 10.1109/MCG.2004.39.

[416] Wu, W., Xu, J., Zeng, H., Zheng, Y., Qu, H., Ni, B., Yuan, M., and Ni, L. M. (2016). TelCoVis: Visual Exploration of Co-occurrence in Urban Human Mobility Based on Telco Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):935–944, doi: 10.1109/TVCG.2015.2467194.

[417] Wurster, C. B. (1961). *KEVIN LYNCH. The Image of the City. Pp. vii, 194. Cambridge, Mass.: Tech nology Press and Harvard University Press, 1960. $5.50*, volume 336. MIT press, doi: 10.1177/000271626133600164.

[418] Yang, D., Rundensteiner, E. A., and Ward, M. O. (2007a). Analysis guided visual exploration of multivariate data. In *VAST IEEE Symposium on Visual Analytics Science and Technology 2007, Proceedings*, pages 83–90. IEEE, doi: 10.1109/VAST.2007.4389000.

[419] Yang, S. N., Schlieski, T., Selmins, B., Cooper, S. C., Doherty, R. A., Corriveau, P. J., and Sheedy, J. E. (2012). Stereoscopic viewing and reported perceived immersion and symptoms. *Optometry and Vision Science*, 89(7):1068–1080, doi: 10.1097/OPX.0b013e31825da430.

[420] Yang, T., Chen, F., Kimber, D., and Vaughan, J. (2007b). Robust people detection and tracking in a multi-camera indoor visual surveillance system. In *Proceedings of the 2007 IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 675–678. IEEE, doi: 10.1109/icme.2007.4284740.

[421] Yang, Y., Cordeil, M., Beyer, J., Dwyer, T., Marriott, K., and Pfister, H. (2020). Embodied navigation in immersive abstract data visualization: Is overview+detail or zooming better for 3D scatterplots? *arXiv*, doi: 10.1109/tvcg.2020.3030427.

[422] Yang, Y., Dwyer, T., Jenny, B., Marriott, K., Cordeil, M., and Chen, H. (2019). Origin-Destination Flow Maps in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):693–703, doi: 10.1109/TVCG.2018.2865192.

[423] Yang, Y., Jenny, B., Dwyer, T., Marriott, K., Chen, H., and Cordeil, M. (2018). Maps and Globes in Virtual Reality. *Computer Graphics Forum*, 37(3):427–438, doi: 10.1111/cgf.13431.

[424] Yeo, W. S., Berger, J., and Lee, Z. (2004). Sonart: A framework for data sonification, visualization and networked multimedia applications. In *ICMC*.

[425] Yu, L., Efstathiou, K., Isenberg, P., and Isenberg, T. (2016). CAST: Effective and Efficient User Interaction for Context-Aware Selection in 3D Particle Clouds. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):886–895, doi: 10.1109/TVCG.2015.2467202.

[426] Zeckzer, D., Gerighausen, D., and Müller, L. (2016). Analyzing Histone Modifications in iPS Cells Using Tiled Binned 3D Scatter Plots. In *2016 Big Data Visual Analytics, BDVA 2016*, pages 1–8. IEEE, doi: 10.1109/BDVA.2016.7787042.

[427] Zelenyuk, A., Imre, D., Cai, Y., Mueller, K., Han, Y., and Imrich, P. (2006). SpectraMiner, an interactive data mining and visualization software for single particle mass spectroscopy: A laboratory test case. *International Journal of Mass Spectrometry*, 258(1-3):58–73, doi: 10.1016/j.ijms.2006.06.015.

[428] Zenner, A., Makhsadov, A., Klingner, S., Liebemann, D., and Krüger, A. (2020). Immersive Process Model Exploration in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2104–2114, doi: 10.1109/TVCG.2020.2973476.

[429] Zhang, S., Demiralp, Ç., Keefe, D. F., DaSilva, M., Laidlaw, D. H., Greenberg, B. D., Basser, P. J., Pierpaoli, C., Chiocca, E. A., and Deisboeck, T. S. (2001). An immersive virtual environment for DT-MRI volume visualization applications: A case study. In *Proceedings of the IEEE Visualization Conference*, pages 437–440, San Diego, CA, USA. Ieee, IEEE, doi: 10.1109/visual.2001.964545.

[430] Zhao, J. and Cheung, S. C. S. (2007). Multi-camera surveillance with visual tagging and generic camera placement. In *2007 1st ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC*, pages 259–266. IEEE, doi: 10.1109/ICDSC.2007.4357532.

[431] Zhou, D., Frémont, V., Quost, B., Dai, Y., and Li, H. (2017). Moving object detection and segmentation in urban environments from a moving platform. *Image and Vision Computing*, 68:76–87, doi: 10.1016/j.imavis.2017.07.006.

[432] Zhou, G., Tan, Z., Cheng, P., and Chen, W. (2004). Modeling and Visualizing 3D Urban Environment via Internet for Urban Planning and Monitoring. In *Structure*, pages 12–23. `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.131.1248`.

[433] Zollmann, S., Poglitsch, C., and Ventura, J. (2016). VISGIS: Dynamic situated visualization for geographic information systems. In *International Conference Image and Vision Computing New Zealand*, volume 0, pages 1–6. IEEE, doi: 10.1109/IVCNZ.2016.7804440.