

Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration

Maxim Spur

▶ **To cite this version:**

Maxim Spur. Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration. Graphics [cs.GR]. Ecole Centrale de Nantes, 2021. English. ⟨NNT : ⟩. ⟨tel-03669606v1⟩

HAL Id: tel-03669606

https://theses.hal.science/tel-03669606v1

Submitted on 16 May 2022 (v1), last revised 31 May 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

THÈSE DE DOCTORAT DE

L'ÉCOLE CENTRALE DE NANTES

ECOLE DOCTORALE N° 602
Sciences pour l'Ingénieur
Spécialité : *Architecture et études urbaines*

Par
Maxim SPUR

Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration

Thèse présentée et soutenue à Nantes, 10/09/2021
Unité de recherche : UMR 1563, Ambiances Architectures Urbanités (AAU)

Rapporteurs avant soutenance :

Arzu COLTEKIN	Professeure,	University of Applied Sciences and Arts Northwestern Switzerland (Suisse)
Christophe HURTER	Professeur des universités,	École Nationale de l'Aviation Civile, Toulouse

Composition du Jury :

Présidente :	Sidonie CHRISTOPHE	Directrice de recherche,	Université Gustave Eiffel, Saint Mande
Dir. de thèse :	Guillaume MOREAU	Professeur,	IMT Atlantique, Brest
Co-dir. de thèse :	Patrick LE CALLET	Professeur des universités,	Université de Nantes
Co-encadrant :	Vincent TOURRE	Maître de conférences,	Ecole Centrale de Nantes

Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration

---

Maxim Spur

September 10, 2021

Acknowledgement

I would to express my warmest gratitude to my directors and supervisors, who made this work and everything achieved through it possible in the first place. You gave me a chance, and your never-ending patience, support, and trust allowed me to see this work through to its end, while teaching me a lot about the intricacies of academic life.

A big thanks to my old friends, and to the fellow doctorants who became my new friends: you supported me throughout with your encouragements, and my progress and the success of this work would be not be imaginable without the inspirations from your unique insights, your help, and the fun we shared during this time.

Of course, I would never be where I am today without my family. Your love means everything, and your unshakeable faith gave me the strength to finish this work.



Contents

I Introduction and Initial Research 1

1 Introduction and Context 3
1.1 Outset and Outline . . . 3
1.2 Context: Urban Design . . . 5

2 State of the Art 9
2.1 Visual Perception . . . 9
2.1.1 Ecological Psychology . . . 9
2.1.2 Affordance Theory . . . 10
2.1.3 Perceptual Processing . . . 11
2.1.4 Gestalt Laws . . . 13
2.1.5 Visual Channels and Preattentive Processing . . . 14
2.1.6 Glyphs and Visual Dimensions . . . 16
2.1.7 Into the Third Dimension . . . 18
2.1.8 Conclusion . . . 22
2.2 Immersion and Analytics . . . 23
2.2.1 Presence and Immersion . . . 23
2.2.2 Visual and Immersive Analytics . . . 25
2.3 Urban Data Visualization . . . 27
2.3.1 Urban Visual Analytics . . . 29
2.3.2 Geospatial Data and Urban Applications . . . 31
2.3.3 Geovisualization Overview . . . 33
2.3.4 Immersive Geovisualization . . . 37
2.3.5 Summarizing the Overviews . . . 39
2.4 Conclusion: A Problem Statement . . . 40

II Hypotheses and Methods 43

3 Immersive Environments: Engaging Space for Spatial Data 45
3.1 Engaging Urbanists? . . . 45
3.2 Virtual Space . . . 47

4 Data Layer Separation: Immersive MCVs 51
4.1 Genesis: Taking the Data Layer Concept Literally . . . 51
4.2 Stacked Layers: Related Work . . . 53
4.3 Map Composition and Comparison . . . 56
4.4 Stacking Layers Vertically as MCVs — in VR . . . 60

5 Data Layer Deformation: Immersive Focus + Context 65
5.1 Genesis: "Into the Layer Stack" . . . 65
5.2 Related Work: Focus and Context . . . 66
5.2.1 Benefits and Problems of 3D City Views . . . 66
5.2.2 Overview+Detail and Focus+Context . . . 67
5.2.3 Techniques: Nonlinear Raytracing and Projection . . . 68
5.2.4 Techniques: Geometry Bending . . . 71
5.2.5 Synthesis and Design Considerations . . . 75
5.3 Inverse Stereographic Projection for Multilayered Focus+Context 77
5.3.1 From Layer Stack to Focus+Context . . . 77
5.3.2 The Sphere — an Ideal Shape? . . . 79
5.3.3 The Stereographic Projection . . . 81
5.3.4 Inverse Stereographic Projection for Focus+Context . . . 82
5.3.5 Terrain Distortions and Flattening . . . 84
5.3.6 Other Projection Surfaces . . . 87
5.3.7 Double Stereographic Projection . . . 90
5.4 Conclusion . . . 91

III Implementation and Evaluation 93

6 Virtually Physical Data Layer Stack 95
6.1 Implementation of the MCV Stack . . . 95
6.1.1 Tools and Data . . . 95
6.1.2 Implementation in Unity 3D . . . 97
6.2 Experiment: Comparing MCVs in VR . . . 100
6.2.1 Use Case in Urbanism: Public Street Lighting . . . 101
6.2.2 Study and Task Design . . . 103
6.2.3 Experiment Setup . . . 104
6.2.4 Experiment Process and Participants . . . 107
6.2.5 Devices, Measurements, and Questionnaires . . . 110
6.3 Results . . . 111
6.3.1 Preferences and Rankings . . . 111
6.3.2 Subgrouping Results . . . 114
6.3.3 Performance and Behavior . . . 115
6.3.4 Feedback and Comfort . . . 121
6.3.5 Scenario Choices and Explanations . . . 123
6.4 Discussion of the Study . . . 129
6.4.1 Limitations . . . 131
6.5 Conclusion . . . 132

7 Virtual Data Sphere for Inverse Stereographic Projection 135
7.1 Implementation: Virtual Data Sphere and Other Surfaces . . . 135
7.1.1 Projection by Shader . . . 135
7.1.2 Projection Parameters . . . 141
7.2 Interaction and Navigation . . . 143
7.2.1 Projection Interaction . . . 143
7.2.2 Navigation . . . 148
7.3 Evaluation, Discussion, and Comparisons . . . 151
7.3.1 Iteration and Informal Evaluation . . . 151
7.3.2 Contributions and Effects . . . 153
7.3.3 Comparison to Existing Focus + Context Methods . . . 157
7.4 Domains for Applications . . . 161
7.4.1 Urban Development . . . 161
7.4.2 Augmented Navigation . . . 163
7.4.3 Planetary Visualization . . . 165
7.5 Conclusion . . . 167

IV Conclusions and Future Work 175

8 Conclusions 177

9 Future Work 181
9.1 Further Evaluation 181
9.1.1 Formal Evaluation of the Inverse Stereographic Projection . . 181
9.1.2 Continued Evaluation of the MCV Stack 182
9.2 Enhancements 183
9.2.1 Inverse Stereographic Projection Enhancements 183
9.2.2 MCV Stack Enhancements 184
9.2.3 Hybrid: Overview+Detail + Focus+Context 185

V Annex and Bibliography 187

10 Annex 189
10.1 Gestalt Laws 189
10.2 MCV Study Introduction 191
10.3 Saccade Polar Plots 193
10.4 Additional MCV Configurations 196
10.5 Shadergraph 198

Bibliography 199

Part I

Introduction and Initial Research

1 Introduction and Context

1.1 Outset and Outline

Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration, just as a collection of the terms itself, requires clarification of almost every word it contains. When is a visualization immersive? What is geospatial urban data, and when is it used — by whom? How do visual analytics work? What are some current difficulties in these fields, and are there opportunities for improvement? Researching recent literature on each topic as well as on combinations among them, a broad introduction is necessary to establish a background and foundation, in order to then go deeper and highlight and critique relevant developments, to discover open research questions, and to finally define more concrete problem statements that could perhaps be solved within the scope of this thesis.

To tackle this problem, a logical order of research was drawn up. Starting with a look at literature on and talking to experts from the field of urban design provided the context, within which further literature could be seen and evaluated for its relevance. Having established that high-level context, the literature research goes back to the basics of visualization. This itself is grounded in visual perception: what are the psychophysical and cognition-based fundamentals that need to be taken into account for effective visualization?

These topics and questions leads to the fields of immersion and presence, as ways to potentially increase the efficacy of visualizations even further, utilizing recent advances in interfaces and interaction methods. It is here that the format and structure transitions from providing a background and context into more of a cursory look at the state of the art — what can be applied inside the previously established context? This part, of course, also touches on the technologies themselves, such as virtual and augmented reality displays, and their applications for visualization purposes, such as visual analytics, defined as the field of utilizing human spatial reasoning of analyzing complex data.

After this groundwork, a look is taken at developments utilizing the above, specifically in support of urban design and exploration, with recent examples from 2D and 3D geovisualizations, as well as a look at the concepts of overview+detail and focus+context, and how they are applied to geospatial and specifically urban data visualization.

Synthesizing this, the outline for the following section on context, the following chapter containing the state of the art, as well as that of the thesis in general is described below.

The second half of this chapter (Context: Urban Design) provides the necessary background with a brief overview and definitions of what urban design is and what makes visualization important for that field. The next chapter (State of the Art) is structured in these four sections:

1. **Visual Perception:** An introduction to the basics of visual perception from different approaches and their implications for visualization.
2. **Immersion and Analytics:** A short discussion of the sense of presence and immersion, and how the field of immersive analytics was recently created from its origins in visual analytics with the help of immersive technologies.
3. **Urban Data Visualization:** An overview of developments in the field, and a critique of recent examples of geospatial urban data visualization, along with concepts from geovisualization that are applicable in this context.
4. **Conclusion: A Problem Statement:** A recapitulation of the state of the art, concluding with a broad problem statement for this thesis.

This first part of the document (section I — **Introduction and Initial Research**) ends with the conclusion of the state of the art and arrives at a first problem statement derived from researching it. The next part (section II — **Hypotheses and Methods**) starts with posing two general hypotheses regarding immersive environments and their uses for spatial data (Immersive Environments: Engaging Space for Spatial Data), and continues with proposing two different methods to investigate these: Data Layer Separation: Immersive MCVs, and Data Layer Deformation: Immersive Focus + Context.

The third part then, as its name implies (section III — **Implementation and Evaluation**), implements and evaluates both methods separately (Virtually Physical Data Layer Stack, and Virtual Data Sphere for Inverse Stereographic Projection), each with their own discussions and conclusions.

Finally, the fourth part (section IV — **Conclusions and Future Work**) concludes by looking back on the hypotheses made earlier (Conclusions) and gives an outlook on Future Work that can come from this. The **Annex** contains some additional materials and illustrations, followed by the bibliography.

1.2 Context: Urban Design

An overview of what can be considered a part of different facets of urban design (UD) can be gathered from a discussion in the book *Introducing Urban Design: Interventions and Responses*, edited by Clara Greed and Marion Roberts [GR14]. Therein, the multitude of equally valid definitions of what can be considered UD is being acknowledged.

According to this book, the differences in said definitions stem mostly from their origin: architects can have a different notion of what UD is about, compared to e. g., traffic engineers or town planners. The latter, for instance, may say that UD is but a small portion of town planning and consists of the aesthetic aspects of it. Traffic engineers, meanwhile, can see UD as a means of traffic calming. Architects can consider UD to be just architecture on a larger scale, and landscape architects can see their own activity as a kind of UD — by designing the landscape in which a city is built, or the landscapes within a city, they are shaping the city itself.

In addition to these observations, the book also provides two more succinct definitions of UD: one from the University of Westminster and the other from Britain's Department of the Environment, both of which are worth being included here verbatim:

> "Urban design is concerned with the physical form of cities, buildings and the space between them. The study of urban design deals with the relationships between the physical form of the city and the social forces which produce it. It focuses, in particular, on the physical character of the public realm but is also concerned with the interaction between public and private development and the resulting impact on urban form."

> "…urban design should be taken to mean the relationship between different buildings; the relationship between buildings and the streets, squares, parks and other spaces which make up the public domain; the nature and quality of the public domain itself; the relationship of one part of a village, town, or city with other parts; and the patterns of movement and activity which are thereby established: in short, the complex relationships between all the elements of the built and unbuilt space."

The book's editors further provide this concise summary: "Each planning decision has a design implication and it is in this implication that the urban designer's sphere of influence lies."

Apart from perceptions and definitions, perhaps a better way to understand an urban designer's work and to differentiate it from others is through a generic example of a

major new urban development project. Let's consider just three spheres of participating actors: Town planners, transport planners, and urban designers. The town planners will be concerned about economic and environmental impacts of the project, such as the effects on residents, jobs provided, potential destruction of natural countryside and agricultural land, and the influence on the whole region. Transport planners will consider the relationship to surroundings from a different point of view. The existing provision of highways is examined along with how additional traffic congestion could be caused by the new development's increased traffic volume, and what the impact on existing road capacity will be. Finally, the urban designers are concerned with the design of the new scheme itself: its layout, the relationships between buildings and space, and its access and circulation patterns. It is important for them to consider how the design will work as a scheme, what its implications are for future growth potential, and also the impact on neighboring areas. Of course, the actual visual appearance is of key importance as well.

In summary, urban designers are said to be players in decision-making processes which shape the built environment. They stand at the nexus of the town's management and urban governance, investment finance and property developers, community politics and environmental sustainability. Lastly, urban designers are particularly concerned with the interaction between people and their urban spaces, and strive to extend the participation of non-professionals in UD.

While that book gives a general overview of the high-level tasks urban designers face, more recent and more focused publications can help us pivot from this to our concrete topic and its underlying basics, which is urban data visualization. Two good examples here are *In search of visualization challenges: The development and implementation of visualization tools for supporting dialogue in urban planning processes*, by Monica Billger et al. [BTW17] and *Visual Analytics in Urban Computing: An Overview*, by Yixian Zheng et al. [Zhe+16].

As can be understood from the urban designer's position at a crossroads between many different professions and their desire to foster broader participation [Hea97; Ama+10], the communication of ideas, concepts, and plans is a very important part of their work, and today that of course relies heavily on information technology [Han07]. That first paper [BTW17] expands on this importance and provides a multitude of examples from research literature that illustrate how visualization tools are or could be helpful in this communication. It then goes into analyzing the discussed literature to find and categorize certain common challenges in urban visualization for dialog purposes. Aside from developing tools that make the dialog process more engaging, the more basic problems of avoiding misinterpretation or finding the best ways to represent data are listed as part of the five main challenges the authors have identified.

The second paper [Zhe+16] meanwhile focuses on urban visualizations not mainly as a tool for discourse, but for helping the urban designers make sense of the increasingly vast amount of data that is becoming available to them. It argues that this data could be very useful in assisting decision-making and helping formulate better plans for future city developments. Since with increasing amounts of data and complexity the understanding of it becomes ever more difficult, this paper also strives in part to collect recent developments from the field of visual analytics as it applies to urban computing. Visual analytics, here, is the subject of visualizing data in a way that aids human experts in understanding it, seeing relevant patterns in it and generally making sense of otherwise non-intuitive data.

Examples from both these papers shall be discussed later as appropriate, but this direct quote from *Visual Analytics in Urban Design* may serve as a bridge to the next topics — visual perception and visualization:

> "…it is widely accepted that human perception plays an essential role in data analysis. With the increasing volume and complexity of urban data, it is unlikely that [computers] can take the place of human beings in a decision-making process in the foreseeable future. It will be more and more important to figure out an efficient way to involve humans in the process of data analysis. Although visualization provides a potential solution, it should not stand alone, but should integrate seamlessly into the applications of diverse domains."

2 State of the Art

2.1 Visual Perception

A discussion of visual analytics and the visualizations it builds on that are so central to the modern urban design process would be incomplete without starting at their foundation in *visual perception*. However, a complete state of the art report on this topic would range from physics and physiology to psychology and philosophy — how are visual signals created and transmitted, in both physical and virtual ways? How are they registered and interpreted by the human visual system, starting with the transmission through the optics of the eye, then the transduction into electrical signals by the retina, to finally be processed and interpreted by the layers of neurons in the visual cortex of the brain — while actually being pre-processed with complex feedforward and feedback loops along the way [Spi99; KBÖ15]? Is perception merely the process between being stimulated and responding accordingly, or is it the process of forming meaningful shapes from simple and often unrelated elements? Do perceptions constitute real aspects of the outside world, are they just parts of one's own mind, or are the perceiver and the environment dynamically and structurally coupled?

2.1.1 Ecological Psychology

In the introduction to his seminal 1979 book *The Ecological Approach to Visual Perception* James Gibson [Gib79] wrote how just concentrating on the physical facts from optics and anatomy is not enough to fully understand vision. Experts in physics and physiology, he writes, appeared to be steadily approaching models or theories of perception that could eventually be validated in experiments. That approach is flawed in his argument, because he has seen how in vision research certain assumptions were made, like comparisons between the eye and an artificial camera. The restraints of studying vision by looking at the cellular level or examining test subjects in constrained settings (especially by limiting the field of view, or by asking to fixate on particular points) can only yield very limited, if useful, understandings, Gibson argues. As the visual system evolved in animals that are free to move around, to see objects from different sides and to shift their attention to new vistas, Gibson opines that only taking those facts into account allows one to develop a real understanding of what he calls *natural vision*. This includes not just the often-studied *snapshot* and *aperture* vision, which both limit the fixation of the eye to

be more still camera-like, but *ambient* and *ambulatory* vision, which allow looking and walking around, respectively.

Gibson does not flatly reject the more physical and physiological insights, but regards them as being tangential to the true question of what visual perception really is. In this final book of his he instead asks the following questions:

> "How do we see the environment around us? How do we see its surfaces, their layout, and their colors and textures? How do we see where we are in the environment? How do we see whether or not we are moving and, if we are, where we are going? How do we see what things are good for? How do we see how to do things, to thread a needle or drive an automobile? Why do things look as they do?"

His proposed answer to those is a theory he calls *Ecological Psychology*, which models the "awareness and activities of individuals in an environment" [McL03]. A more thorough discussion is far beyond this scope and not necessary here, but one central idea from Gibson's final work in dealing with the above issues has had deep implications in vision research and, more recently and critically, in visualization and interaction: his *Affordance Theory*.

2.1.2 Affordance Theory

Coined by Gibson, the term "affordance" can be used to describe physical properties of objects and environments that animals can perceive to be actionable, based on their understanding of the world in general and of objects or components thereof in particular. A prime example is the handle on a teapot; an obvious affordance for holding it. Generalized away from vision or physical interaction and applied to psychology, it includes all possible transactions between individuals and their environment. Perception to Gibson, then, is not a bottom-up, atomistic affair as was the theoretical consensus before him: it is not points of light that are perceived and then processed and assembled into a greater understanding of the environment. It is the possibilities for action in the environment that are directly perceived, and not indirectly perceived through individual sensory evidence. After all, it makes evolutionary sense to develop a visual system that is good at perceiving possibilities for action.

Colin Ware's influential 1999 book *Information Visualization: Perception for Design* [War99], while being critical of Gibson's rejection of the mechanics underlying vision, is highly supportive of his theory in its application to visualization, particularly interactive ones. Even if the theory of "direct perception of environmental affordances" may be problematic because of the inherently indirect nature of visualizations, the affordance

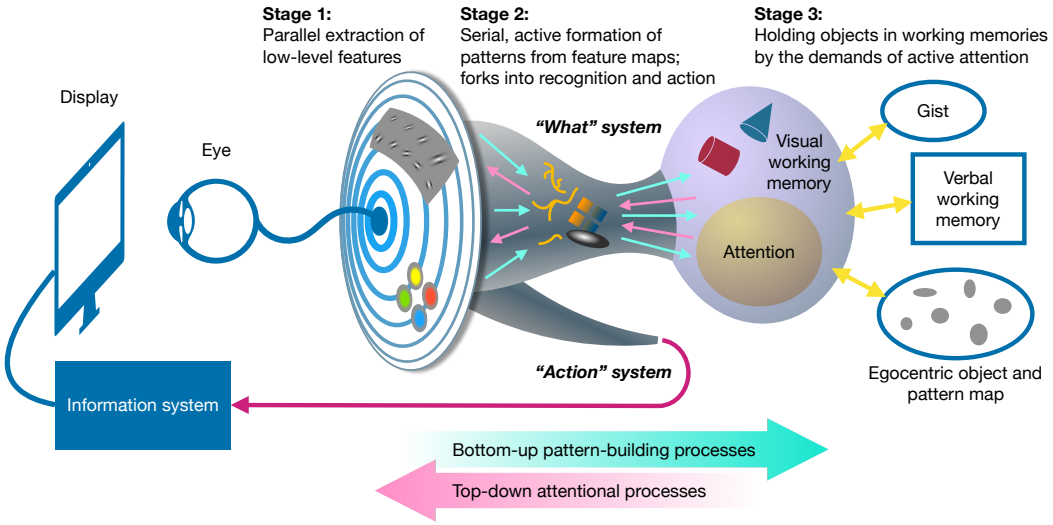


Fig. 2.1: A three-stage model of visual information processing *(adapted from [War13]).*

concept "can be extremely useful from a design perspective" for shaping interactions. It seems desirable to build interfaces "that beg to be operated in appropriate and useful ways," and components should be made "perceptually evident" in the ways they are able to work together and to be acted on. Yet, as Ware adds, the affordance theory only goes so far by helping create good interaction designs for visualization. For a more complete "science of visualization," a more inclusive and holistic view of perception, which is not as dismissive of physical facts, is necessary, Ware argues. Rejecting the "importance of understanding visual mechanisms would be to reject a tremendous proportion of vision research as irrelevant," especially since modern display technologies are based on precisely understanding those same mechanisms. This is where the model of *Perceptual Processing* comes in.

2.1.3 Perceptual Processing

This model, which Ware calls a "simplified information-processing model of human visual perception" and first introduced in his 1999 book, consists of three connected stages responsible for different parts of visual perception. A conceptual drawing of the stages, their elements, and how they are connected is shown in Figure 2.1. While being very broad, this model provides a conceptual framework that is useful in discussing the individual systems that make up visual perception. The greatest part of Ware's book is in fact structurally built and ordered around its three stages [War13].

First, in an automatic and highly parallel process, visual information is condensed into features by neurons on the retina and the primary visual cortex, which are selectively tuned to detect certain parameters like colors and orientations of edges, as well as movement.

In the second stage, those features are put together into patterns, such as continuous contours and regions of the same textures. Interestingly, this pattern-processing stage is forked into two systems: the "*what* system" and the "*action* system." The first perceives and identifies objects, while the second leads to brain areas involved in controlling physical actions, thus creating a feedback loop used e. g., for locomotion and closer examination. This split was documented in [GM92] through physical and behavioral evidence, and the *two-visual-system theory* was later expanded on in [MGV95]. Quoting from the first source: "the set of object descriptions that permit identification and recognition may be computed independently of the set of descriptions that allow an observer to shape the hand appropriately to up an object."

The third stage is where up to about five objects are held in visual working memory. Those are formed not only from the patterns discovered by the previous stage, but also come from long-term memories that are related to a task. Ware presents the example of finding a route on a map by searching for highways between two cities, represented by connected red contours and visual symbols, respectively.

At this point it is important to note that the flow of information does not only go in one direction between these stages — this is only the case when a completely new image appears in the visual field. In all other cases, the whole system is being tuned from top to bottom to enhance and consolidate what appears on the lower stages based on expectations and usefulness.

These interactions between the three systems are what makes up visual *attention* and involve the whole visual system. The bundles of neurons responsible for detecting features in the first stage can be tuned to be more or less sensitive, depending on what features are being sought. The second stage is directed by the last one to seek out patterns by both forming them from the first stage's output and by actively steering eye movement to different possible patterns or by following multiple patterns, which can form complete scenes and objects.

Ware emphasizes the importance of these interactions in what he calls "one of the more radical ideas" in his book:

> "... the effects of attention can be propagated outside of the brain into the world through cleverly designed interactive visualizations that cause information we are interested in to be highlighted on the screen."

The interactivity part will be a major topic of this thesis partly for the above reason, but this will be discussed further down. For now, it is clear how pattern perception plays a central role in the visual perception system. To better understand what those visual

patterns are and how they can be categorized, it is helpful to take a look at a concept that originally stems from psychology, but has since been refined and partly re-defined: the *Gestalt Laws*.

2.1.4 Gestalt Laws

The Gestalt Laws, which were formulated as a result of the first serious attempt at understanding pattern perception by the "Gestalt School of psychology" in early 20th century Germany [Kof35], were a strong influence on Gibson and are to this day a valuable way to look at basic perceptual phenomena, as Ware states. "Gestalt" can be translated from the original German as "shape," "form," "pattern," or similar, but lacks a direct equivalent in English. Basically, the Gestalt laws describe how and which patterns are easily perceived by humans, both visual and temporal ones. Examples are elements that are organized or easily grouped by their spatial proximity, their visual similarity (e. g., shape, color or texture), their connectedness to each other, and a common fate, i. e., if they are moving similarly or otherwise changing in a similar way over time.

Comparable to how the affordance theory does not have a sound basis in the way perception works on a physiological level — by design, in fact — the neurological mechanisms originally proposed by the Gestalt School researchers have mostly not been found to be accurate. Still, as with the affordance theory and perhaps even more so, the Gestalt Laws have found great application in visualization and interaction design; they "easily translate into a set of design principles for information displays," according to Ware.

Ironically, these "Laws" are not a fixed set of universal principles. For instance, Palmer and Rock [PR94] argued for a new law, that of *connectedness*, which was not discussed by the original Gestalt psychologists. An example of the power of connectedness as a grouping principle are node-link diagrams, used to illustrate graphs since at least the 13th century [Mor14]. On the other hand, the quite general principle of "similarity" has largely been superseded by the "channel theory and the concepts of integral and separable dimensions," as Ware explains.

In his books, Ware discusses a total of eight Gestalt laws and how they all contribute to the *figure-ground effects* that were investigated by the original Gestalt psychologists, i. e., how these laws contribute to creating a *figure*, which lies above a clearly separable *ground*. Examples adapted from [War13] are shown in figures 10.1 to 10.6 in the Annex (section 10.1). Other authors have compiled different lists of Gestalt laws, according to their needs [CDT02].

A more recent and very thorough overview of the state of the art in Gestalt Psychology can be found in [Wag+12], which includes additions such as the principles of *Synchrony*

and *Common region*, and which further elaborates on the differences of those grouping principles between static and dynamic patterns. The authors further make a distinction between perceptual grouping and figure-ground organization, and list a multitude of recently recognized principles.

2.1.5 Visual Channels and Preattentive Processing

These philosophical, psychological, and high-level understandings are important to keep in mind, but an even greater source of applicable insights has come from *physiological* research into the anatomy of the visual system. The current understanding of it explains how each step of vision functions on a neurological level, from the different light receptors and their locations on the retina, over to how low-level processing is done by the layers of neurons right behind those receptors, to how signals are processed by the two visual areas of the cortex V1 and V2. Aside from explaining how visual illusions like Mach bands or the Cornsweet effect are created, those understandings show the kinds of patterns that the visual cortex processes.

In an article by Livingstone and Hubel [LH88], physiological and anatomical observations show how there are "several separate and independent subdivisions that analyze different aspects of the same retinal image." Cells have been found that are specifically receptive to discerning properties like orientation and size, color, local motion, and local stereoscopic depth. It follows that those properties are perceived independently and therefore form separate visual channels that do not interfere with one another. One example of the implications for visualization is that single symbols can express multiple dimensions of information, encoded in different channels like color or orientation. Precisely understanding the kinds of patterns the visual cortex responds to allows for the creation of efficient visualizations, as Ware argues.

It is important to note the difference between the above research, which falls under neuropsychology and psychophysics. The latter forms theories about perception by conducting direct studies of human responses to physically defined stimuli. Visual distinctiveness as studied through experiments with human subjects resulted in a list of features that are "preattentively processed" [TS85], i. e., they are instantly recognizable and separable at first glance. These include line orientation, width and length, size, curvature, blur, numerosity, color, and hue, among others. While some of these features match those discovered by neuropsychological research or can be arrived at through combination of the visual channels mentioned above, some others are not, like the distinction between straight and curved lines. It may be that receptors on a cellular level have not yet been found for those missing effects, or that different kinds of mechanisms are at work here.

A mostly still current state of the art of preattentive processing is presented in an article by Healey and Enns [HE12] [1]. In addition to listing currently known visual features that are instantly recognized, sometimes multiple theories on how they work are provided, with information on how those inform recent work on memory and attention. Another important part of this publication is its overview on how those findings are currently being applied to visualization and graphics, and also which areas are in need of more research. An example is the influence of aesthetics and engagement on attention and visual understanding, which were still only in the hypothesis stage at the time of publication.

Important to note here is that not all visual channels provide this so-called "popout" effect (another name for "preattentive processing," as often used by the visualization community), and not all features that are preattentively processed are directly linked to a neuropsychological structure. "Visual channel" in this context refers to the term used by the visualization community, meaning a "way to control the appearance of marks, independent of the dimensionality of the geometric primitive;" a *mark* here being "a basic graphical element in an image,"] and a "geometric primitive object" that is classified by the number of spatial dimensions required to display it (e. g., points, lines, areas, etc.). The effectiveness of these visual channels in providing popout depends on "both the channel itself and how different the target item is from its surroundings," as Tamara Munzner explains in her 2014 book "Visualization Analysis and Design" [Mun14].

Aside from their ability to provide popout, visual channels can be categorized and ordered in a number of ways. Munzner first outlines two principles: *expressiveness* and *effectiveness*. The first one dictates that "the visual encoding should express all of, and only, the information in the dataset attributes." This basically divides visual channels into *identity* and *magnitude* channels. and assigns them to categorical and ordered data, respectively. Importantly, Munzner states that "unordered data should not be shown in a way that perceptually implies an ordering that does not exist," and vice versa.

The *effectiveness principle*, on the other hand, dictates that "the most important attributes should be encoded with the most effective channels in order to be most noticeable," leaving the question open of what actually makes one visual channel more effective than another. Munzner provides a ranking of effectiveness for both identity and magnitude channels, and a discussion of the reasons for it based on a channel's *accuracy, discriminability, separability*, and their ability to provide visual popout and *perceptual groupings*.

Still, those lists and rankings are not definitive by any means and Munzner herself poses a number of related research questions in some of her most recent talks [Mun17]:

[1] A more recent addition to the list of preattentive techniques is based on dichoptic presentation — presenting different stimuli to each eye, e. g., by presenting a target only to one eye [KK19]. Since this does not require altering any visual properties of that target, it is well-suited to visualization, and particularly useful in augmented or virtual reality devices which can directly control what is shown to each individual eye [Kre+20].

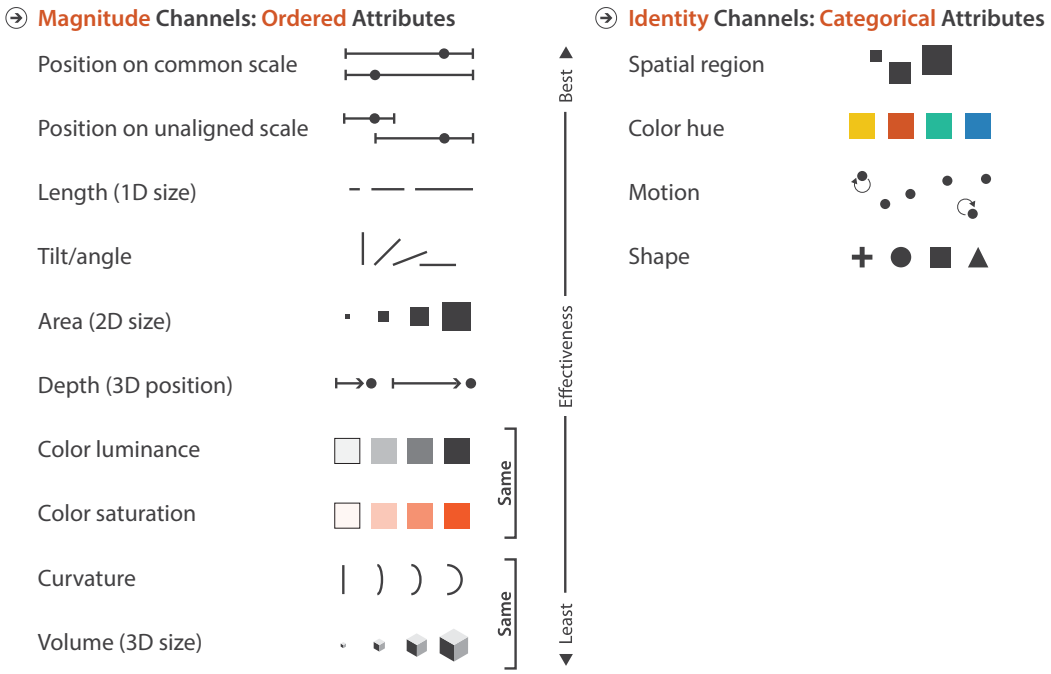


Fig. 2.2: Ranking of visual channels on their effectiveness, as proposed by Tamara Munzner [Mun14].

> Is this the correct order of channel rankings? What are the rough equivalence classes? What are the metric distances between them? What are the trade-offs between them? What is the capacity of channels? How many discriminable steps/bins/levels are there in each channel?

Even more questions arise from using combinations of channels, in addition to how using multiple channels at once influences the above questions:

> How, and how much do the channels interfere with each other, and to which extent? What is the effect on the number of discriminable steps? What about n-way interactions between channels for n up to at least 5–6?

To better understand these interactions between channels, it is worth taking a look at Garner's theory of integral and separable dimensions.

2.1.6 Glyphs and Visual Dimensions

A *glyph* in visualization is a "graphical object designed to convey multiple data values," as Ware explains. It is in designing these glyphs where visual channels can clash and the interactions between them particularly matter. In his 1974 article *The Processing of Information and Structure* [Gar74], Wendell Garner laid out his concept of which display attributes (visual channels) will be perceived independently from another. This concept

divides visual dimensions (yet another synonym for attributes or channels) into *integral* and *separable* dimensions, based on how pairs of these are perceived. Through a number of experiments (restricted and speeded classification tasks) it was determined which pairs of a glyph's attributes are perceived holistically (*integral dimension*), and which pairs allow independent judgments on both attributes more easily (*separable dimensions*). An example is a colored ellipse: its width and height (two integral dimensions) form one holistic shape, from which its color is a separable dimension. The separable dimensions are also said to be *analytically processed*; implying that the choice of integral vs. separable dimensions for visualizing data depends on how the viewer is expected to react and make judgments — on the combination of dimensions, or separately on each variable.

Colin Ware notes in his book that while there is extensive evidence for this dichotomy and real use for it as a design guideline, it still is very simplistic as a concept. Since there is "always some interference between different data values presented using different graphical attributes of a single visual object," the dimension pairs effectively lie on a continuum between being separable and integral. For some dimensions, like stereoscopic depth, there is as of yet too little evidence available to categorize their pairings with other dimensions. Others, like color, are dividable into sub-dimensions, like the "degree of redness" of a color, which complicates the creation of a definitive list. Tamara Munzner, for instance, argues in her aforementioned book for four visual channels inside the "color" channel: *luminance, saturation, hue and transparency*. Another problematic aspect of the integral–separable dichotomy is the fact that the experimental paradigms for determining these classifications are all related to interactions between pairs of variables only, and very little work exists on three or more variables, as Ware states and Munzner reiterates in her open research questions.

Apart from the above attribute of integrability/separability, visual dimensions can be said to be *monotonic*, i. e., whether they are able to express relations such as "greater than or lesser than" in one way or another. Ware argues that such qualities are size, brightness and vertical height, while others, like hue and line orientation are not. Furthermore, some monotonic qualities express a natural direction that is clearly understood to mean an increase, like the vertical height above the ground plane and the vividness of a color. Some others, like the horizontal position, while being monotonic, can be understood and ordered differently depending on cultural background, e. g., if the writing system is left-to-right, although recent research suggests a general rightward bias [Wal15].

A final attribute of dimensions from Ware's book is how many discrete, resolvable steps are encodable in them, particularly for rapid or preattentive processing. While Ware just lists rough guidelines, e. g., four steps in orientation and size, eight steps in color, etc., Munzner states this outright as an open research question.

2.1.7 Into the Third Dimension

All above deliberations concerned themselves primarily with planar 2D visualizations. Figure 2.2 does list *depth*, meaning the 3D position of objects, far below e. g., *length*. This informal ranking of *effectiveness* encompasses multiple dimensions, an important one of which is *accuracy*. Different sensory channels have been examined through psychophysical experimentation to determine this accuracy: can we accurately judge objective changes in intensity?

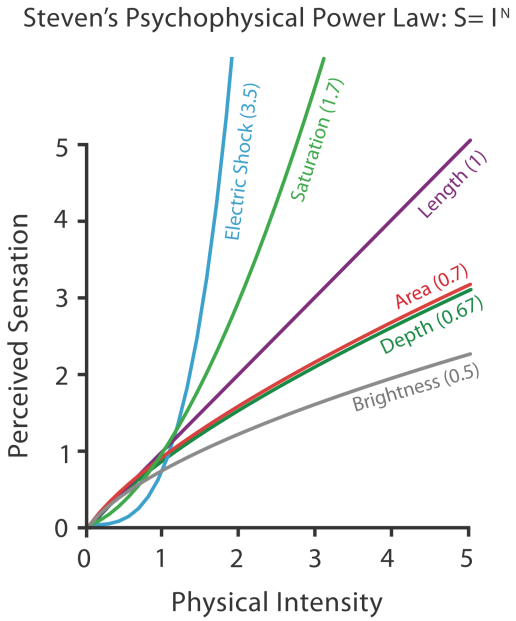


Fig. 2.3: Stevens' Power Law, showing the relations between physical intensity of a stimulus and its perceived intensity.

Stevens' Power Law [Ste17] (Figure 2.3) is a measure that shows how certain channels compress or magnify these changes: most attributes are actually not perceived linearly. A doubling in e. g., area does not appear to be double as large, while small changes in color saturation appear more intense than they actually are. Length is an exception, and ranks therefore very high in accuracy — its intensity maps linearly to our perception. Important to note here is that this length strictly means the *planar spatial position*, i. e., distance of co-planar features, where the plane itself is seen as such. This plane does not need to be exactly at the image plane of the viewer, i. e., fully perpendicular to the viewing angle, as the visual system can compensate for a certain degree of tilt and still perceive planar images correctly [VGB05].

From a pure standpoint of conveying information using visual channels, 3D does not add much, as the distance in 3D space of features is not nearly as accurately perceived as distances on 2D planes (c.f. Figure 2.3: Depth vs. Length), and actually introduces a number of downsides in terms of effective visualization. As we are visually much more

familiar with the 3D world we inhabit than the abstract 2D planes used for illustrations, we expect certain cues in depictions of objects in 3D, which must be carefully considered, lest they become inaccurate at best or misleading at worst. Perspective is an obvious one: correct perspective projection of 3D to a 2D screen looks more natural, but results in foreshortening, which distorts the relationships in size of objects at different distances. We also expect three-dimensional objects to be shaded to some extent, which can interfere with using the color channels for information. Occlusion is also very helpful in establishing a depth order, but can by its very nature hide objects and information.

These and other expectations are what actually enable us to notice depth at all, as it is not something we can perceive directly, but rather have to interpret from *depth* or *distance cues*. These can be divided into three categories: monocular static, monocular dynamic, and binocular cues [War13]. Monocular static cues are ones that we can extract from a non-moving 2D image using only one eye, like the mentioned shading and perspective cues, as well as gradients in size and texture (which compress in the distance), focus depth, relative sizes, and occlusions — these also often interplay and depend on each other. The number of these cues that are available for this most accessible category (as it requires nothing more than printed images) are what enables us to perceive depth, and therefore three-dimensional shapes to varying degrees of accuracy in e. g., sketches, drawings and photographs. In fact, cues from this category are more important for depth perception and involve less effort to depict than those from the other two categories.

Binocular cues result from the distance between our two eyes. Eye convergence, or *vergence*, is basically a result of sensing the angle of the eyes when they fixate on an object, and using that information to triangulate its distance. This is in practice limited to arm-length distances, after which *stereoscopic depth* takes over in importance. That cue results from the (often slight) differences between the images both eyes receive from the slightly different angles they see objects from — *retinal disparity*. These two cues can also be classified as primary (physiological) and secondary (pictorial) [HS05], or as extraretinal and retinal cues, respectively; both interact and contribute to distance perception [MT99], and were already discussed by Gibson in relation to a "psychophysical theory of perception" of the visual world [Gib50].

Egocentric distance plays a large role in which cues are predominant for depth estimation, and can also be used to divide the environment into different perceptual spaces: *personal space* (limited to 2 m), *action space* (up to a radius of 30 m), and *vista space* (beyond about 30 m) [CV95]. For shorter distances (in the personal and closer action space), the availability of accommodation, convergence, and retinal disparity results in "remarkably accurate" distance estimation, barring "consistent and well-researched minor distortions" [DH09]. Beyond those closer distances, gross over- and underestimations of egocentric distances can frequently occur, depending on the available cues and conditions. The *action* space here takes on its other meaning, as an observer could potentially walk up

to an object: this kind of measurement would be an *action cue*. For distances in vista space, pictorial cues, such as *aerial perspective* (visibility reduction through fog and other atmospheric particles) [Ros67], *texture gradient* (compression of the density of texture elements) [HU88], and *geographical* and *optical slant* (the relation of the line of sight to the surface, or inclination relative to the surface) [GC52] are most significant [DH09] — these are the *monocular static cues* mentioned above.

Naturally, the perception of especially the physiological cues can be hindered or completely prevented by certain impairments, leading to loss of *distance stereoacuity* [HF05], up to the point of *stereoblindness* [Ric70]. The prevalence of stereoblindness is not clear, but estimates range between 1 and 30% of the population [ZKF03].

For the non-stereoblind majority of the population, stereoscopic displays that offer binocular cues "can provide a particularly compelling sense of a 3D virtual space, and for a few tasks they can be extremely useful," writes Ware [War13]. Visualizing these cues by definition requires a separate image for each eye, which can be provided even on printed or otherwise static media through (auto-)stereograms or anaglyphs. For dynamic visualization on active displays, the most typical systems involve either glasses that filter out images intended for the other eye, such as polarized lenses or active shutter glasses and a specialized monitor or projection system, or head-mounted displays (HMDs) that place a separate monitor close to each eye.

Finally, there are the monocular dynamic cues, which are also referred to as *structure-from-motion* cues. The only two kinds of it are *motion parallax,* where different objects moving at in the same direction and at the same speed appear to have a gradient in their velocities depending on their distances to the viewer, and *kinetic depth,* where the 3D shape of an object becomes apparent through rotation relative to the viewer — we generally assume objects to be rigid, and thus the apparent changes in shape is interpreted as movement in space. Of course, visualization of these cues requires a step up from static displays; an animated screen becomes necessary.

Allowing animated or even stereoscopic visualizations introduces trade-offs that arise from technological limitations. Depending on the complexity of the scene or animation and the capability of the computing hardware, the refresh or frame rate of the display may fall below what is necessary to convey fluid movement, or impede on the perception of the system's responsiveness, which has been determined to be important to user performance [Wat+98]. Stereoscopic displays — or rather those providing *binocular presentation* [Vis14] — pose a number of complex challenges that need to be addressed well in order to avoid visual discomfort [KL01].

If the route of HMDs is chosen, *physical* discomfort from the devices' weight can also become a problem, as well as the phenomenon of *cybersickness*, which arises if a discon-

nect between the visually perceived movement, and the actual movement as reported by the vestibular system [LaV00] is noticed. This issue and its precise mechanisms are still under intense investigation [RO16; WKB19], as are techniques [AQ20] and guidelines [Por+17] to mitigate it.

A recent and more complete discussion of depth cues and their perception, with a mapping of these to existing or proposed display technologies can be found in section 2.2. of one of Marriott's contributions [Mar+18a] to the 2018 book on *Immersive Analytics* [Mar+18c]. In it, the nowadays almost commonplace *head-mounted binocular displays* (referred in this work as just HMDs) are deemed as capable of conveying all depth cues except *accommodation* — the "effects of dynamic physiological changes in the shape of each eye and the consequences on the retinal image," or the blurring caused by focusing at different depths.

While this capability already enables a high degree of depth perception, the lack of accommodation (due to these devices using fixed lenses, and therefore fixed focal distances) can lead to misestimations of distances, especially in AR headsets through which the real environment is simultaneously visible [SES18]. The problem becomes clear when there are virtual and real objects at the same egocentric distance, but are not equally in focus, or the other way around. Advanced hardware prototypes using flexible lenses has been tested that can provide even these [Dun+17], but recent work has shown that another, already commercially available kind of display technology which directly projects images on the retina [Lin+17] — rendering *all* content in focus — can also significantly reduce these misestimations [Pei+20]. The *vergence-accommodation conflict* now only occurs outside the area a user focuses on, and is thus less jarring. Other factors impairing depth estimation in HMDs are the often severe limitations in a device's field of view, and even its weight — both seeing improvement with new hardware generations [BYB18].

Depth perception — up to this point just talked about as the perception of egocentric distance from an object — can be even further complicated when considering exocentric distance estimation. The actual distance between two objects on a plane that is perpendicular to the line of sight is arrived at from combining the quite accurate length estimation between them [VGB05], and the estimation of both objects' distance to the observer. As discussed above, the accuracy of this varies greatly with distance. Any other arrangements complicate these estimations even further, as does displaying them virtual reality devices, which have been shown to cause a clear anisotropy of distance perception [Pei+19b], and augmented reality, where combinations between real and virtual objects are asymmetrically misestimated [Pei+19a]. On the other hand, the ease with which virtual environments can be displayed and manipulated allows a user to navigate within them, or displace and rotate them until more favorable views are achieved that reduce misestimations — VR/AR has the potential to bring anything into the *action space*.

For abstract data visualization, these challenges of using 3D depth might not seem to warrant the effort for such a relatively imprecise visual channel, and certain visualization textbooks actively discourage such use of "unjustified 3D" [Mun14]. Where it is indispensable, though, is "when the user's task involves shape understanding of inherently three-dimensional structures," which is precisely the case in this work. Kevin Lynch already stressed the importance of representing the shape of the city as accurately as possible in his seminal work *The Image of the City* [Lyn60], and this extends to (virtual) 3D models being beneficial for understanding urban issues [Pol+18].

2.1.8 Conclusion

Visual perception has been approached from angles as diverse as philosophical reasoning, psychological experiments and neurophysiological studies. While oftentimes at odds with each other, these disciplines have contributed to our current, well-rounded understanding of how the human visual system works and how we can design visualizations take the most advantage of its immense processing power. We understand how this system evolved to where it is now, how it functions on a cellular and structural level, and we can form predictive concepts of its workings that have been confirmed and refined by psychophysical experimentation.

Investigation in this field is far from being concluded, however, with a large amount of ongoing research addressing open questions. Some additional examples that were not mentioned above are the perception of symmetry [PA18] and ensemble formation [CC18]. Contributing to advancing the field are recent developments in neuroscience, as well as technological developments that allow more intricate psychophysical experimentation, such as ever more accurate and less invasive eye tracking devices and more capable techniques for analyzing the gathered experimental data [Kur+14].

The current guidelines for visualization based on these understandings, redefined and refined over decades should be kept in mind for any use case that delivers information visually. Specifically in our case, some of the most important "data" to be visualized is actually the shape of the city. If this 3D shape, including elevation and building geometry is of importance, as it is increasingly becoming with the vertical development of cities and the increasing importance being placed on perceived density [Fis18] and visibility of the sky [YPL07] or other features [Ort+16] in urbanism, the often disregarded visual channel of depth takes on an integral role for appropriate visualization.

In order to lessen interference with the other visual channels that encode additional information than just the physical shape of cities, be it thematic geospatial maps [HCM13], trajectories [GAM15] and other spatiotemporal data [Dor+18], atmospheric data [SPG12], or zoning and planning constraints [Bra+16], we can aim to maximize the availability

of depth cues by using HMD technology. Not only does it provide stereoscopic vision, but the tracking of the user's movement in 3D space allows for the much more powerful structure-from-motion cues without requiring direct interaction with the visualization. Lastly, they can be more helpful than most tools in creating a sense of *presence* and *immersion* into the virtual scene and the data, which is also very helpful in the task of visual analytics — both topics for the next section.

2.2 Immersion and Analytics

These technological advances in virtual environments are approaching an "almost full coverage of the user's perceptional system," and have driven the advent of the "Spatial Presence Theory" [Har+15]. *Presence* — the subjective psychological experience of being in a virtual or remote space — and *immersion* — the objective characteristics of the technology used to present the virtual space [Dwy+18] — go hand in hand, and influence one another in various ways[SW97]; they are also themselves heavily dependent on the technical devices used to achieve these states. Coming back to Gibson's *Ecological Approach to Visual Perception* [Gib79], he implicitly laments the lack of these concepts in much of daily activity:

> We modern, civilised, indoors adults are so accustomed to looking at a page or a picture, or through a window, that we often lose the feeling of being *surrounded* by the environment, our sense of the *ambient* array of light [...] We live boxed up lives.

As stated in the foreword to the 2018 book titled *Immersive Analytics* [Mar+18c], the same problem befalls even the quite recent discipline of *Visual Analytics*. As "the science of analytical reasoning facilitated by interactive visual interfaces" [CT05], most of its practical applications and tools still use standard desktop computers as their platform. The field of *immersive analytics* is now aiming to integrate technologies on the *Extended Reality* (XR) spectrum — Virtual, Mixed, and Augmented Reality — with visual analytics to hopefully develop more engaging and efficient analysis and decision support tools and imbue them with the otherwise missing immersion.

2.2.1 Presence and Immersion

Display and interaction technologies play a large part in determining how *immersive* a system is, in the sense of providing a sensory experience close enough to the real world to convey a feeling of being *present* in the virtual world. The introduction to immersive analytics [Dwy+18] presents an overview of research into these aspects since the inception of the idea of *telepresence* in 1980 [Min80].

Similar to how using more visual channels increases the effectiveness of visualizations, engaging additional sensory channels increases the *extensiveness* of the system — adding e. g. spatialized audio is an obvious way to heighten a sense of spatial presence. More complicated setups that have studied the effect of additional channels found a significant increase in such presence when providing even more stimuli, such as thermal, wind, and olfactory sensations [Ran+18]. The *inclusiveness* of a system is determined by how much these sensations block out those from the real world.

Vividness, or the degree of the virtual environment's fidelity, appears differentiated in its effect on presence: pure visual realism did not impact it as much (if at all) as more lower-level factors, such as stereopsis [BHB99], geometric field-of-view, latency, and frame rate [SS05]. While the field-of-view is still limited in most current consumer-grade HMDs, it is seeing constant improvement, and their frame-rate is actually in most cases already superior to the standard 60 Hz refresh rate of common desktop monitors. This fluidity of motion also works together with the inherent spatial tracking technology to mirror a user's motions in the real world: the more seamless is appears, the better the *proprioceptive matching* is [SW97]. The *plausibility* of the system is increased by the inclusion of dynamic shadows [SUC95], and generally by how realistically the virtual world itself and objects in it behave [Sla09].

Gibson's Affordance Theory and his Ecological Psychology approach also find their way here as a feature of presence: the *autonomous action* [Sme; FH98]. From that perspective, actions and interactions, or rather what can be done in the virtual world take precedence over visual fidelity — the system's *affordances* for action are more important than photo-realism [Sch10]. In extreme cases, even a text-based environment can lead to a higher sense of presence than a fully-featured 3D world, if the user can realize their goals in the former but not in the latter [Riv09].

The same can be true for *immersion* when used as a term for an individual experience, rather than as a technical descriptor [WS98]. [2] Much of the research into immersion as a psychological phenomenon has been focused on computer games — its high importance there was emphasized, but it was not always clearly defined as a term [BC04]. Game immersion has been categorized into three stages that build on each other: *engagement*, *engrossment*, and finally *total immersion*; ranging in the amount of time and effort players are willing to invest, and even how much emotionally involved they become in the game. As with the affordances' effect on presence, game immersion can often depend more on narrative elements than on graphical fidelity. These factors are however related, and Slater's sensory immersion [SW97] does play a role in engagement [EM05; CCN14].

[2] Slater's definition and reiterations [SW97; Sla99; Sla09; Sla18] of the term *immersion* as the "objective level of sensory fidelity a VR system provides" has for some time now been the common way to use it within the VR community, especially when contrasting it to *presence*, referring to "a user's subjective psychological response to a VR system" [BM07], but it is nonetheless interesting to explore another use of the term in this section.

Another concept linked to this perceived immersion in this sense is *flow* [NC14], which was introduced even before telepresence as "the enjoyable subjective state that people experience when they are completely involved in something to the point of forgetting time, fatigue and everything but the activity itself." Helping to achieve this state, or a similar high engagement level of *cognitive absorption* is of course highly desirable, and is now being seen as the next step in designing software interfaces that go beyond "just" usability [OT08; Att+11]. Here again, immersion in Slater's sense — sensory fidelity [SW97] — is seen as one method to further this goal, and this is what distinguishes immersive analytics from pure visual analytics in the field of visualization.

2.2.2 Visual and Immersive Analytics

Visualization, as a distinct discipline that builds on computers' ability to interactively render physical and abstract data to visual displays, began to develop itself as soon as computing devices started gaining graphical capabilities. After the advent of the first oscilloscope-based displays in the 1950s, it did not take long for *computer graphics* to become recognized as a sub-field of computer science — at the latest, with the introduction of the ACM SIGGRAPH conference series in 1974. Principles from graphic design and perception [Ber67; CM84] and statistics [Tuk77] were quickly adapted to work with computers, and (physical) scientific data became easier to visualize as 3D capabilities grew, especially once efficient hardware dedicated for that purpose emerged [Cla82]. A workshop on "Visualization in Scientific Computing" in 1987 [MCC87] formally established that discipline, with the first IEEE Visualization conference following in 1990. An intense focus on visualizing abstract information [TGB90; CRM91] — mapping any kind of data spatially — led to further specialization, with the IEEE Symposium on Information Visualization in 1995 presenting itself as a contrast to pure scientific visualization, and eventually becoming its own conference in 2007.

Technological progress also lead to an exponential growth in data, especially with the advent of communication over the internet. The sheer amount of data to analyze, along with certain security concerns emerging in the early 2000s, led to an urgency to develop effective analysis methods that could cope with this volume and complexity. With the new National Visualization and Analytics Center, a new science concerned with "analytical reasoning facilitated by interactive visual interfaces" was born — *Visual Analytics* [CT05]. This original definition, as outlined by the 19 initial high-level recommendations for that field, did not specify actual devices for the display of and interaction with visual analysis systems, but at this time, the information visualization community was rather skeptical of even using 3D graphics, so the focus of visual analytics defaulted to traditional desktop environments [Dwy+18].

The progress made in 3D graphics made it possible to develop applications that took advantage of the third dimension to display data. Examples such as cone trees for hierarchies [RMC91], and attempts to take advantage of "spatial memory" [Rob+98] seemed like obvious steps forward up until the late 1990s, but subsequent studies started to curb this enthusiasm by not finding any benefits over more traditional 2D alternatives for these specific examples [CM00; CM01]. The eventual proliferation of "gratuitous" use of 3D for even simple bar and pie charts [SP08] makes the subsequent backlash against it among the community and their calls for caution [Mun14] understandable.

During this phase of enthusiasm for 3D, early immersive technology also found use for visualization — a natural fit for scientific applications with 3D data, but pioneering work also utilized VR for abstract data [CRM91]. The earliest VR and AR systems can even be traced back to Sutherland's work in the 1960s [Sut68], but the technical limitations were obviously quite severe at this time. When the first interactive VR systems that were considered "practical" started appearing more than two decades later [Cru+92], they continued to require complex and expensive hardware setups while still delivering subpar visual fidelity compared to contemporary traditional display methods.

Decades of steady technological advances finally led to affordable, easy-to-use, and relatively high-fidelity commercial VR products by a number of manufacturers, starting with the first Oculus Rift prototypes in 2012–2013, the higher-grade HTC VIVE in 2015–2016, and even AR headsets with the Microsoft Hololens in 2016. These devices have since evolved into different product lines, targeting different market segments ranging from industrial applications to home consumers, with some even aimed squarely at the gaming sector, e. g. Sony's PlayStation VR and Valve's Index, and others attempting to integrate them into social networks — Facebook's Oculus Quest, as part of that company's foray into the creation of a *Metaverse*.

These suddenly very accessible XR devices alleviated many issues that previously plagued them and held back their utility for visualization. With years of research now behind interaction with virtual content (the ACM Symposium on Spatial Interaction was started in 2013), the case to embrace the new paradigms XR is capable of has been made at a workshop called "Death of the Desktop: Envisioning Visualization without Desktop Computing" at the IEEE VIS 2014. The term *Immersive Analytics* was coined one year later [Cha+15], with a number of workshops following every year at various IEEE conferences, and finally the *Big Data Visual and Immersive Analytics* conference in 2018 explicitly calling for submissions in that field.

Immersive analytics stands to deliver on the original goals Thomas and Cook set out for visual analytics [CT05], and potentially engage the senses more fully, as well as engage more potential users, be used in new contexts away from desktop computing, and enable new kinds of collaborative work. Furthermore, research on user experience has shown

how it is strongly linked with engagement and productivity [HT06] — improving one has a positive effect on the others. **This focus on engagement in immersive analytics, and the effect XR has on it in general is a strong motivating factor to explore these technologies and principles and how they can be applied for an improved user experience when exploring and analyzing urban data.**

The book *Immersive Analytics* by Marriot et. al. [Mar+18c] tells the story of the field in more detail [Dwy+18], calls to "reconsider the value of 3D for information visualization" [Mar+18a], and explores the various facets of presence and immersion as they relate to effective visual analytics [McC+18], as well as XR's technical capabilities for interaction [Büs+18]. A new design framework for immersive analytics that extends the classic "What, Why, and How" in visualization [BM13] is presented [Mar+18b], as well as one for understanding *situated analytics* — "the use of data representations organised in relation to relevant objects, places, and persons in the physical world for the purpose of understanding, sense-making, and decision-making" [Tho+18]. Its final chapter investigates applications of immersive analytics to simulated built environments [CMK18] — the main concern of this thesis, and the topic for the next section.

2.3 Urban Data Visualization

Visualizations of urban data, in its most frequently encountered form, are maps: "symbolic representations of the world drawn to an accurate scale, using a set of uniform symbols" [Day94]. Before the drive for increased accuracy for purposes of recording land transaction and the military — made possible by the introduction of accurate surveying techniques during the Renaissance — led to the now familiar plan projection maps, cities were usually depicted from an aerial view, showing both their geometric form and street layouts. In these drawings, accuracy did not play the major role and "could be sacrificed in order to ensure that all the finest buildings and monuments were shown and the artistic and symbolic qualities of the work were at least as important as its objective status" [Ell87].

Alan Day posited in 1994 [Day94] that though the usefulness of today's easy-to-read maps — facilitated through a high degree of abstraction — for "navigation, planning new developments, or mapping spatially related data" is clear, the power of this paradigm "tends to blind us to other forms of representation." At this time, Geographic Information Systems (GIS) gave maps "intelligence" by linking them with databases and adding analytic tools, but were still limited to two dimensions: heights were abstracted to contours, and the three-dimensional form of cities was excluded from representation as a "complexity of the real world" that stood in the way of "accuracy, consistency, and portability." Even in a time when a single 3D view of a Computer-Aided Design (CAD) model of a city could

take dozens of hours to render, he already argued for bringing CAD and GIS together: computer modelling could achieve both the accuracy of modern maps while realistically representing the form of cities. The map should in his view be placed as "just one level of representation within a much richer environment."

From the title of it (*From map to model: the development of an urban information system*) and the arguments and conclusions within, Day's article could be seen as advocating for city models to supersede maps, at least for certain use cases. Urban models can provide an accurate context to view new building designs and proposals in, to make sure they fit within their surroundings. Most importantly, they "provide a representation of the physical form of the city which everyone can understand," and can represent changes to it in an intelligible way to nonspecialists. "Democratic" urban development relies on informed choices, made ideally not only by small minorities of experts, but by residents and their representatives. Digital 3D models are thus a very welcome technique to facilitate debate by making information more accessible: easier to understand than maps, and much more flexible (let alone portable) than traditional physical scale models [THF06].

The way he contrasts models to maps, however, actually also speaks to the strengths of the latter. The "high degree of abstraction" maps employ to "reduce the amount of information to manageable proportions" allows them to be read easily. These abstractions can be helpful to"concentrate the mind on specific issues," especially to trained professionals. His article can thus be used to argue that both maps and models have their place in urban visualization, depending on use (creation, analysis, communication, etc.) and user (designer, stakeholder, layperson, etc.), even if the previously discussed '90s enthusiasm for 3D (section 2.2.2) clearly shines through his writing.

Since this thesis explicitly does not focus on any one use case or user group, instead aiming to generally investigate immersive approaches to geospatial urban data visualization, maps and models are both considered. Since Day's recommendations in 1994, significant progress has been made in using maps (be they 2-, 2.5-, or 3-dimensional), digital models, and immersive technologies to visualize cities. This section highlights a number of relevant and recent work in these areas, including from geovisualization not specific to urban environments, but nevertheless very applicable there. Of special interest here are examples that also incorporate approaches from visual (and immersive) analytics, or at least explore how immersive technologies and interaction techniques can contribute to the recently formulated discipline of *Urban Analytics*[SSF17]: the "application of quantitative, computational, design and visual methods to the spatial and morphological structure of cities and regions," [Bat17] and "the core set of tools employed to deal with problems of big data, urban simulation, and geodemographics" [Bat19].

2.3.1 Urban Visual Analytics

A relatively recent, and very comprehensive overview of this kind already exists in *Visual Analytics in Urban Computing: An Overview* [Zhe+16]. In contrast to Batty's very broad definition of "urban analytics," The more technical work by Zheng et. al. focuses on the application of visual analytics on "urban computing" [Zhe+14] — "a multi-disciplinary research field, where computer science meets conventional city-related areas, such as civil engineering, transportation, economics, energy engineering and environmental science, which usually leads to complex analytical tasks for real world applications." Many such tasks are not easily automated, at least not fully, and they thus still rely on human perceptual ability and domain knowledge — visual analytics aims to bridge this gap by combining "automated analysis with interactive visualization for effective understanding, reasoning and decision making."

This overview [Zhe+16] categorizes works on this "urban visual analytics" into two classes: *data exploration and pattern interpretation*, and *visual learning*. These two classes build on each other: visual analytics tools are "designed for better insights of different types of urban data," which aid human exploration and interpretation, but to create these tools, a certain amount of automated analysis or at least structuring of the data needs to exist. Sophisticated mechanisms like data mining algorithms or machine intelligence can provide the necessary output, but need to be fine-tuned or trained to be useful — visual learning, as a special case of active learning [Pai+14] is an effective method to accomplish this. That topic of course exceeds the scope of this thesis, and the examples from the first category are very specific solutions for particular problems, but provide a great overview over the many different types of urban data, and how different visual channels are employed in their visualization.

Examples in that overview [Zhe+16] include singular visualizations of certain data on one map (e. g. density maps of trajectory data [Sch+11]), but many others are centered around systems that link multiple visualizations together; either connecting abstract charts with maps (e. g. combining parallel plots with 3D visualizations [Fer+15]), or different layers or slices of data on the same map (e. g. atmospheric, energy, and transportation data in one city [KSR12]), or show different views, scales, and interpretations of similar data (taxi trip visualizations [Fer+13b]). A similar overview by Luo and MacEachren [LM14a] focuses less on urban environments, and more on "geo-social" data, such as social networks and other human behavior. Batty also presented a less in-depth outline of "visual analytics for urban design" [BH14], highlighting how new visualization tools, especially those that utilize real-time data, are "beginning to enrich the field of urban design."

A more recent article went at this task from the other way, looking at urban data analysis from a visual analytics perspective [Dor+18]. Less of a state-of-the-art report than the

other works above, it condenses the challenges visual analytics of spatiotemporal urban datasets faces (visualization, data analysis, and data management) and uses a small number of examples of their own work, including the TaxiVis [Fer+13b] and Urbane systems [Fer+15] to show how they can be addressed. The authors took on the "visualization challenge" in their examples by designing novel clustering techniques [Fer+13a] that reduce the heavy cluttering that is inherent in the dataset they visualized: large amounts of origin-destination data from urban taxi trips. They also included their interactive visual querying method [Fer+13b] under the umbrella of "visualization challenges," as it makes visual exploration of such data easier and more effective than the traditional SQL queries they compare it to. Their answers to the "data analysis challenge" include a computational topology-based technique that captures "spatiotemporal dynamics across a city over multiple temporal resolutions" from various datasets [Mir+16], as many urban phenomena are not sufficiently described by single sources of data. The "urban pulses" this system identifies are placed on maps, with charts visualizing their properties. Finally, they developed a "time lattice" data structure [Mir+18b] to address the challenge of managing and querying large amounts of time series data, which "enables sub-second responses to queries over time series having even a billion points, while supporting interactive updates of new data."

The article stresses how the sheer volume, complexity, and heterogeneity of urban data makes the traditional practice of "confirmatory data analysis" — driven by domain experts formulating hypotheses based on theories or anecdotes, to be analyzed by data scientists, and finally confirmed by experts again — difficult to scale and adapt to modern concerns. Interactive and fast visualization systems that incorporate data analysis — visual analytics — facilitate exploration and insight generation on a previously impossible level, especially with well-designed low-latency systems [LH14]. The increasing verticality of urban landscapes is brought up in the final part of the article as what necessitates a "mov[e] beyond flatland" to support data analysis and exploration in 3D, like in the use case of shadow accrual on streets and buildings [Mir+18a].

These overviews and others like it stress the importance of visual analytics in geospatial domains in general, and urban environments specifically, and demonstrate its progress there; *urban data visualization*, when using interactive techniques to generate insights and aid decision-making, becomes *urban visual analytics*. Many of their examples stay within 2D maps [Sch+11], but some do display 3D models of a city's buildings [Fer+15] or use the third dimension for other data [KSR12], or to display spatiotemporal variations in a *space-time cube* [KL04; Tom+12]. What is common throughout is the wealth of information that these visualizations contain, often utilizing many of the available visual channels, or, conversely, splitting off into multiple views. The following subsections will take a closer look at a few such systems, and also take into consideration those made more generally for geovisualization, as many advances in that broader field can certainly apply to the more specific urban environment.

2.3.2 Geospatial Data and Urban Applications

The very recent book titled *Manual of Digital Earth* [GGA20] contains two more overviews of relevant advances. Published twenty years after Al Gore's call in 1999 for a "Digital Earth" [Gor99] — a "multi-resolution, three-dimensional representation of the planet, into which we can embed vast quantities of geo-referenced data," tackling the problem of more and more georeferenced data being accumulated without ever making sense of it or it contributing to any analysis — it answers that call in a multitude of practical and theoretical ways. Partly inspired by Buckminster Fuller's *Operating Manual for Spaceship Earth* [FS69] which precedes it by five decades, the book "presents a systematic analysis of the theories, methods, and technical systems of Digital Earth" and summarizes developments in disciplines including earth science and the data acquisition, processing, and visualization necessary for it, its applications in climate change research, global development goals, and "digital cities," breaks down Digital Earth developments across different regions, and even ventures into the ethics of such systems and how they can be used in education.

Two of its chapters are of particular interest to this thesis: *Digital City: An Urban Perspective on Digital Earth* [JS20], and *Geospatial Information Visualization and Extended Reality Displays* [Çöl+20]. The former begins by explaining the emergence of "Data Cities" — the latest term to describe what has over the years been called "fractal cities" [BL94], "cities of bits" [Mit96], "intelligent environments" [Dro97], and most commonly, "smart cities" [GKS92] — which was greatly accelerated with the advent and proliferation of satellite and networking technologies, allowing unprecedented navigation, communication, monitoring, and imaging, which in term allowed for the sophisticated modelling and simulation of urban environments and "urban dynamics" [For89] possible today — a "new science of cities" [Bat13].

The imminent threats posed by the climate crisis to cities — and populated areas in general — became clearer through these advances in Earth observation, and many national and international organizations are stepping up to the dual challenges of preventing and mitigating the worst effects, including rising sea levels and extreme weather events. Already, some of the world's most powerful computing centers are being employed or even specifically built to model and simulate complex environmental systems, and yet field experts such as Bob Bishop [Bis20] are warning that "quantities of data — mostly unstructured data — are growing far faster than global computing power." Novel computing systems may be key in keeping up with the growth of input and demand for actionable output, such as quantum computers or neuromorphic chip architectures — both already in various stages of development and even use [MV17; Sen+18]. Beyond using new hardware that imitates the human brain for more efficient processing, artificial intelligence (AI) is mentioned in this overview as having the potential of not just providing e. g. better

and faster classification of sensor data, but to directly assist in planning and governing tasks, or even take decisions itself — already cautioned against expecting too much from it by Batty [Bat18], and explored in more detail in later work by other researchers [Yig+20; KH20] that venture to assess AI's possible impact on cities in the future. A final (and more optimistic) biomimetic idea mentioned in this overview is the call for an automated "Digital Earth Nervous System" [DeL+10] to "provide decision makers with improved alerting mechanisms for crisis prediction and situational awareness."

At this time, such all-encompassing systems are far from being universally implemented. Singapore is mentioned here as one of the most "geospatially advanced" examples, thanks to its "extraordinarily integrated government," which has enabled the creation of its *Geospatial Master Plan* [Auta] that aims to "maximise the impact of geospatial information and technology for the economy and the wider society," as well as the *Virtual Singapore* project [Autb] – a "3D digital replica" of the city that "can be used in simulations and virtual tests of new solutions to urban planning problems." Already, systems have been proposed that integrate this Virtual Singapore database for e. g. modeling microclimate conditions and predicting energy demand, from building to district level [Ign+19], or that augment its sensor-based real-time data with visual crowd-sourced information [HK20]. But these current or even planned systems are still far from approaching the autonomous nervous systems proposed above, and mostly "just" serve to provide more up-to-date, or finer-grained information to the stakeholders and decision makers, often as a part of urban visual analytics.

The examples and case studies this overview presents are less focused on their visualization aspects (if those are even present), and more on the "global potential" such "advanced urban environmental simulations" might have — the overview is more concerned with the data and its application. They are nevertheless interesting to mention, not least as indications of direction as to what actually needs to be visualized, or what is relevant to the field. Before listing their case studies, the overview goes briefly into recent advances of GIS, BIM, and CAD systems, and their recent integration with additional data and simulation tools like *City Engine* [Sch+15] (article not from the overview). It also recaps developments in immersive technologies for use with these systems.

One such case study outlines a number of projects done by the MIT Media Lab, including *Umi*, an urban modeling interface that simulates building energy use, daylight availability, and how friendly an area is to walking, and visualizing the results over a 3D model [Rei+13]. Also from the Media Lab, its City Science Group's *CityScope* project is presented: a "data-driven interactive simulation tool for urban design," which uses a tangible interface made up from LEGO bricks [Alo+18]. Any interaction with the physical UI components, and most importantly with the bricks and tiles that represent roads, parks, and buildings and define the urban shape gets passed immediately to the system's "computational analysis unit," which calculates a number of *urban performance indicators* such as density,

diversity, and even social well-being. These and other indicators are displayed on a dashboard, and heatmaps are projected directly on the physical model.

A report by the "datatecture" studio *300.000 km/s* from Barcelona is mentioned in the overview, which builds on their *atNight* project, focusing on nocturnal landscapes, or *nightscapes*, as a challenge in urban design [SM15]. Their 2017 report, addressed at the Barcelona city council's *Municipal Institute of Urban Landscape and Quality of Life* used data on the city's lighting regulations, satellite images of light pollution, and combined it with mobility data, business locations and activity hot spots to provide analysis and visualizations. Finally, the overview looks at the *Greening Greater Bendigo* project in central Victoria, Australia; the city's "first urban forest strategy" [Gre]. It is relying on multispectral Earth observation imagery from ESA's Sentinel satellites, stacking time-series images into spatio-temporal "data cubes" to "monitor seasonal and area-specific changes in vegetation throughout the city." This clarification in the city's vegetation cover reveals potential risks in new building developments for the greenery.

The authors of that overview summarize it by emphasizing the importance of integrating "area-specific modeling of natural environmental systems with current best practices in building information modeling and city information modeling" — digital cities need to incorporate Earth observation data to "accelerate delivery of Fuller's visions for a '4D Air-Ocean World Town Plan' and efficient management of resources on Spaceship Earth." While the authors also call for the "Digital Earth nervous system" to become "self-aware and be able to obtain and respond more automatically to unprecedented quantities of environmental information—far too much to be processed by humans," we can strive to aid human analysis through better visualization until such time may arrive, utilizing advances in user interfaces and turning the abundance of data into an advantage. This is precisely the focus of the Manual of Digital Earth's [GGA20] seventh chapter [Çöl+20] — less centered on the urban environment and more on geovisualization in general, though nevertheless containing examples specifically visualizing cities: "the Digital Earth framework is relevant to many application areas, and one of the foremost uses of the framework is in the domain of urban science." Since urban visualization could be understood as subset of geovisualization, if a very specialized one, taking a broader look at geovisualization as presented here might prove edifying.

2.3.3 Geovisualization Overview

That chapter [Çöl+20] is a review that takes a wide approach to summarize "the current state of the art in geovisualization and extended reality." Similar to the previously discussed chapter [JS20], it begins by explains the concept of Digital Earth, but then delves much deeper into the visualization of geospatial information, very much from a human-centric perspective based in an understanding of cognition, perception, and interaction,

before discussing immersive technologies and their use in geovisualization. This focus already becomes clear in their introduction: while acknowledging the effort required in attempting to digitally replicate real-world objects even at the microscale [3], and the much greater challenges constructing a "digital twin" of entire cities, let alone a "true all containing Digital Earth," they echo Al Gore's sentiment from his original proposal for it [Gor99], in that the greatest difficulty actually lies in *making sense* of the information any such digital model would contain — this is where geovisualization is key.

The authors provide an in-depth overview of geovisualization, which itself begins with definitions of the term, and statements of its purpose. Geovisualization here is differentiated from cartography, which can be seen as the process of the *mapmaker* communicating with the *map user*, with its product being e. g. a map or other kind of visual display. The communication aspect still exists in geovisualization, but the emphasis there lies in the process of "information exploration and sensemaking," and much less on a finished "product," unless an understandings and insights can be seen as such. In this process, "scientists and other experts design and use 'visual geospatial displays to explore data, and through that exploration to generate hypotheses, develop problem solutions and construct knowledge' [Kra03] about a geographic location" or phenomenon. Another description of geovisualization is given from the International Cartographic Association's *Commission on Visualization and Virtual Environments:* it is a "theory, methods and tools for visual exploration, analysis, synthesis, and presentation of geospatial data" [MK01], with the goal of connection people, maps, and process, "leading to enlightenment, thought, decision making and information satisfaction" [DMK05].

Having presented these definitions of geovisualization and what its goals are, the authors shift to their understanding of how to accomplish these goals; how to make "useful and useable geovisualization" [Fuh+05] that can utilize our "ability to think spatially and infer spatiotemporal knowledge from a visualization:" it needs to "support users' tasks and needs, and enable users to interact with the data." They describe in more detail the three areas that are key in this undertaking: **data**, the complexity and quantity of which needs to be tamed and "synthesized in meaningful ways;" **representation** of that data within a visualization environment, and the **user interaction** that facilitates exploration of this data in visual displays through "intuitive and innovative ways."

These areas clearly interact with each other, and there is often a complex interplay between the advances and challenges in each. As a result of the "data deluge" [BHS09] there is now much more **data** to work with, but also a lot of different kinds of data. "Covisualization" of very heterogeneous data is one such challenge: it arose from the *advancement*

[3] "Microscale" in this context — a geographic scope — can refer to individual objects: ranging from small installations, wind turbines, and up to and including entire bridges. In e. g. an EPA proposal for mobile emission modeling [US 01], traffic corridors and intersections are considered microscale, with the other levels being the mesoscale, and the macroscale; referring to "specific roadway links and/or analysis zones," and "analysis over a broad regional area (e. g., county, urban area, state, nation)," respectively.

in Earth observation and other spatial data gathering capabilities, as well as from new methods of "spatializing" previously aspatial data [SB97]. Visually integrating such data in a legible way can be a challenge in designing graphical **representations**, e. g. by interpolating between aerial imagery, vector data, and a digital terrain model in order to control perceptions of photo-realism [HC17]. Another way to covisualize would be to use *multiple and coordinated views* [Rob07]; linking heterogeneous data visualization views together, instead of attempting to integrate them into one view, as "the visual integration of such heterogeneous data requires the careful design of graphical representations to preserve the legibility of the data" — an effort the multiple view approach can forgo. Such methods could be further enhanced by using immersive technologies, for which even greater care is necessary in designing appropriate **interactions** — a degree of "seamlessness" is central to preserving their immersiveness.

The authors mention another intersection between these key areas: nonuniform displays, such as *focus + context* [RM93], or "context-adaptive" [Pin+12] displays — a recent survey brings these and similar techniques that vary scale, detail, generalization level, rendering styles, etc. across a single display space together as "interactive lenses" [Tom+17]. Heterogeneous data is represented differently, depending on how the user interacts with it. And yet again, these advances pose even more challenges, as summed up by the title of a 2012 editorial in *The Cartographic journal*: "More maps, more users, more devices means more cartographic challenges" [GF12] — the development of technologies and solutions, the multiplication of data, its uses and users, only make it harder to "[integrate] these knowledge sets to inform representation and tool design."

Noting some areas of interest for later (**heterogeneous data visualization** and **nonuniform displays**), some other works referenced in the introduction of the three key areas are worth mentioning. Bertin's *Semiology of Graphics* [Ber67] is cited here as providing a "conceptual framework, the visual variables, that allows for us to consider the graphical representation of geospatial information at a fundamental level." These *visual variables*, as originally proposed, were position, size, shape, color value, color hue, orientation, and texture. Notably, this proposed set of variables was based on Bertin's reasoning and intuition as a cartographer, as opposed to being "evidence-based" but have found hold as a core concept not only as a tool for mapmaking, and recently also captured the interest of information visualization [Car03]. [4] They have since been extended to include dynamic [DiB+92] and three-dimensional variables [Rau+15], bringing that set closer to alignment with Munzner's visual channels and even the preattentive visual features [HE12] from research in visual perception.

[4] It is also interesting to note that in information visualization expert Tamara Munzner's work [Mun14] these are called visual *channels*, rather than *variables*, even when referencing Bertin's work, and her list of these channels is not quite the same: compare Figure 2.2 in the above section on visual perception. As she herself states, though: the term *channel* "is popular in the [visualization] literature and is not meant to imply any particular theory about the underlying mechanisms of human visual perception," and is merely a synonym for *variable*, or even *attribute*, *dimension*, *feature*, or *carrier*.

The authors stress the importance of understanding the intended users when designing visualizations, which can go as deep as approaching cognition itself as it relates to visual-spatial displays [Heg11; ODF15], to adhering to user-centered design methods [LD11] and embracing user studies [ÇFL10; Rot+17] to e. g. "evaluate strengths of new methods of interacting with visualizations," especially since geovisualization has turned map *readers* into map *users*. This understanding is crucial, especially when the "designer is not the user," [RTC15] as is often the case.

Similar to how *urban visual analytics* emerged from visual analytics that is specific to cities, the "science of analytical reasoning with spatial information using interactive visual interfaces" has been referred to as *geovisual analytics* since 2007 [And+07], with goes beyond geovisualization's primary aim of supporting data exploration to "coupling computational methods with interactive visualization techniques" to support complex reasoning processes. Though more complex a field, developments there can certainly be insightful to better "simple" geovisualization, and, of course, urban data visualization. The authors dedicate a large portion of their overview to progress in visualizing spatiotemporal data, specifically the visualization of movement. Examples are trajectory maps: the space-time cube concept [Häg70] has been employed in various ways for single and multiple trajectories [Kra08], and also encoding additional attributes of the way [Tom+12], or as just one facet of a flexible, immersive system used to explore very large numbers of trajectories [Hur+18b]. Remaining in "simple" 2D maps, statistical and data mining techniques can be used to reveal clusters from large sets of trajectories depending on desired characteristics, and compared against each other [AA10]. Here the authors point out different methods available for the kind visual comparisons systems like this facilitate, such as **switching between** options, **superpositioning** them, or **juxtaposing** them in *small multiples*, and these methods being better than using animations [Rob+08].

The authors continue their overview of spatiotemporal visualization by investigating depictions of movement without concrete trajectories, where only the origin and destination is known, such as *origin-destination maps* [WSD11] and *flow maps* [Jen+18]. Usually visualized as lines and curves on maps, these can quickly become overwhelmingly cluttered and occlude the map and themselves. To combat this, computer graphics techniques such as *edge bundling* have been applied to reduce that clutter and increase the visibility of significant patterns [HV09]. Another possibility is to take advantage of the third dimension, especially if used in immersive environments: Yang et al. have implemented and evaluated flow maps on "3D globes" and other global map projections for virtual reality, with a clear advantage showing over 2D representations [Yan+18a].

2.3.4 Immersive Geovisualization

At this point, the overview shifts to its part that is most relevant to this thesis: immersive technologies, and their potential to transform geospatial sciences. The authors begin this part by briefly outlining the "essential concepts for immersive technologies:" the original taxonomy of XR technologies (the Virtuality Continuum, [MK94]), and a quick introduction to the terms *presence* and *immersion*, as they are used in this context — see section 2.2.1. The continuum of *immersiveness* has been proposed as a way to order 3D visualization types — from 3D charts, over photorealistic models, to displaying virtual environments in a CAVE — *and* the technologies necessary to display them. A second axis, the continuum of *visual realism* is added to create a "rough classification of 3D visualization" [ÇLZ16] that is particularly relevant for geovisualization, where e. g. terrain and buildings can be represented in numerous ways, from using only contours and wireframes, to being shaded in realistic ways or for visualizing overlaid data. The importance of considering interaction is reiterated, as classical input methods from the desktop environment (mouse and keyboard) are usually not practical in immersive environments where the users are standing or moving around.

The authors first review augmented reality in its use for geovisualization. As the technology for practical HMDs is still riddled with challenges, many of the examples they list are prototype *mobile augmented reality* (MAR) applications [Tsa+16] running on smartphones or tablet devices, which e. g. superimpose point-based information on the real environment [VDV14], or project 3D models of landscapes in an indoor setting [CB17]. An interesting exception is an AR "sandbox," [ABP18] which projects elevation and environmental data on a literal box of sand that is being scanned by depth cameras and interacts in real time to changes made by users, very similar to the brick-based CityScope mentioned before [Alo+18]. While currently limited by its physical requirements, a similar system can be imagined fully in head-mounted AR, where instead of manipulating sand or bricks, users' hand and finger motions can be tracked with enough precision to emulate the same interactions — Microsoft's recently released HoloLens 2 is in fact already approaching these capabilities.

Another exception is a proposal for a versatile in-situ AR application, specifically aimed at supporting urban design — a process "dedicated to plan and preview a future version of the urban spaces" [Dev+18]. AR is envisioned here to mostly display life-sized 3D models: transplanting a photogrammetrically captured alley into a corridor, showing the result of a construction site as seen from an office window, and even digitally removing existing buildings on a real street and (partially) replacing them with planned models. The last example in particular places high demands on the ability of an AR device to position and orient itself in the real world, and accurately overlay 3D models while taking lighting conditions into account, as well as handling occlusion and visibility: if a real

structure is removed, what is behind needs to be replicated and blended in with the surroundings. Current see-through AR HMDs are disadvantaged in this last regard by only being able to *add* light on their otherwise transparent displays, and not fully occlude what is seen of the real environment. Strictly *mediated* reality displays, which record the real environment and display it back after being "mixed" with virtual components, such as smartphones, do not suffer this problem. *Mixed Reality* (MR), as conceptualized in the virtuality continuum [MK94] broadly encompasses both these and the additive, see-through AR HMD technologies — MR is defined as anything between the real world, and a fully virtual one. It also includes another subcategory: *augmented virtuality*, in which real objects are integrated into an otherwise fully virtual environment. Recently, "multimediated reality," among other concepts has been proposed [Man+18], but current research appears to be more focused on the handheld, and with the latest advances in commercially available AR HMDs, the see-through AR paradigms. The authors did explicitly "not expound on MR" other than these in their chapter, as any information on AR they present is equally applicable there, as is most of the information of their next part: Virtual Reality.

In their introduction to VR, the authors draw the distinction between virtual *reality* and virtual *environments*: the latter "encapsulates the range of things one can do," and is basically a part of the former, which is rather seen as a display type, or a "communication medium" [SC02] — the virtual environment is the *virtual world*, one of the four criteria that make for a VR display; the others being immersion, interactivity, and sensory feedback. *Virtual Geographic Environments* (VGEs), since their first conceptualization in 2001 [LG01], have over that time evolved to become described as being "generated by computers and related technologies that users can use to experience and recognize complex geographic systems and further conduct comprehensive geographic analyses, through equipped functions, including multichannel human-computer interactions (HCIs), distributed geographic modeling and simulations, and network geo-collaborations" [CL18], and have as such "attracted considerable attention in the geographic information science research community over the last few decades." It is important to note that VGEs are by no means only used as content for virtual or augmented reality displays, they stand apart as a concept from particular visualization types. By their nature, they are however a natural fit for use in immersive environments, and the geospatial community "could enrich and enliven the augmented reality experience by fusing VGE/3D simulations (as well as their results) as immersive content for augmented reality platforms that blend with everyday real experiences of living and exploring in urban environments" [Tor15].

Being similar to the "digital twin" and "digital Earth" concepts, VGEs can mirror not just the visible geographic surface of a real environment, but additional associated layers of data. As such, a VGE can go "*beyond* reality, as it ideally enables its user to visually perceive invisible or difficult-to-see phenomena in the real world, and explore them inside the virtual world" [Çöl+20]. When including analytic capabilities, a VGE "can be

superior to the real world for analysts," if a realistic experience is desired in their use cases. This combination of GIS with immersive visualization environments becomes very similar to the premise of immersive analytics in the way that it can provide "novel ways to combine human cognitive abilities with what machines have to offer" — going "beyond virtual reality" by providing virtual environments with "geographic processes, phenomena and interactive mechanisms at different scales" that VGEs are built upon [CL18]. A quite complex VGE system has been implemented by Lin et. al., simulating and analyzing air pollution at different geographic scales [LCL13], and demonstrating all four components, or "subenvironments" necessary for a complete VGE: data, modeling and simulation, the interactive component, and the collaborative environment [Lin+13].

The authors finish this chapter on the "current state of the art of topics related to visualization in the context of Digital Earth" with outlining *dashboards*: practical platforms that e. g. package geographical information in a suitable way to be presented to policy makers, or to provide citizens and visitors with access to a "rich tapestry of open data feeds." More accurately defined as "graphic user interfaces which comprise a combination of information and geographical visualization methods for creating metrics, benchmarks, and indicators to assist in monitoring and decision-making" [PL17], these tools are seeing an increasing development and use as the number of challenges that need urgent addressing grows (e. g. "climate change, food security, an aging population, and rapid urbanization"), and relevant big and open data becomes available, or in fact overabundant. Typically, data is presented in multiple windows (graphs, charts, scatterplots, maps, etc.), often linked with each other to enable interactive visual analysis in the form of *coordinated multiple views* [Wid+15]. Various dimensions of e. g. cities or can be explored, from the scale of individual traffic cameras, to traffic congestion and public transport performance, over to air quality and weather data over a whole region [PLJ17]. Other than giving an overview over this kind of simulated and real-time sensor data, similar examples exist for deeper statistical and geospatial analysis, such as housing affordability [PTR17].

2.3.5 Summarizing the Overviews

In summary, this chapter [Çöl+20] from the Manual of Digital Earth [GGA20] discusses a similar problem as the one dealing with *Digital Cities* [JS20]: an often overwhelming amount of geospatial data, and the urgent need to make sense of it. While the latter mostly concerns itself with detailing the kinds of data relevant to urban environments, its sources and applications, the former is providing an overview of means and methods to visualize it, with a focus on utilizing immersive technologies, citing a large number of recent works on many facets of that subject matter. Despite having the broad scope of "geospatial visualization," much of that chapter is applicable to the urban domain, if not even specifically about it. Both chapters name the large **scale** and **heterogeneity** of geospatial data as challenges for sensemaking — the next step for the Digital Earth after obtaining

all this data — with multiple and coordinated views, as well as focus+context and other non-uniform views being presented as possible solutions. Immersive environments are proposed for the advantages in geospatial exploration and analysis they can provide; one cited reference going as far as showing experimentally how they can also be instrumental in "enabling experts in a field [linked to urban projects] to discover new technologies" to their advantage [DMP16].

The overviews presented above [LM14a; Zhe+16; Dor+18; JS20; Çöl+20] come from as diverse perspectives as geography, cartography, social and urban sciences, visual analytics, and geovisualization, to name a few. Each offer their point of view on why visual sensemaking is important in the geospatial (or specifically urban) domain, what challenges lie therein, and which recent developments look promising to advance it. A few common themes regarding urban data visualization can be observed throughout these overviews:

- The importance of 3D for data sources and visualization media,
- the amount and complexity of data that needs to be understood, and
- the potential of immersive environments and technologies for facilitating visual exploration and analysis.

Seeing how these same conclusions can be drawn from those different overviews reinforces their validity.

2.4 Conclusion: A Problem Statement

This chapter's *raisons d'être* are multiple. Its first goal is to clarify the concepts that make up the starting point of this thesis: Context: Urban Design, Visual Perception, Immersion and Analytics, and Urban Data Visualization. What emerges is that these can be explored in a coherent way, where each part and builds upon or even requires another: urban design relies on being able to perceive the city in many ways and be immersed in it. Visual perception and immersion are key not just in getting a "sense of place" in a real or virtual city, but also in performing analytical tasks effectively, which themselves require appropriate visualization. Visualizations of *spatial* urban data can be seen as a part or a specialization of geovisualization, where immersive technologies are increasingly finding traction. The latest overview of such applications [Çöl+20] itself touches on many of the topics discussed above, e. g. visual perception: by understanding the cognition and vision of users of geovisualization, a *user-centered design* approach can be taken.

By touching on a number of works from these diverse, yet related fields, the second goal of this chapter can be realized: discovering common problems and possible approaches to address them. This is most essential, as from the outset, this thesis was quite vaguely

defined (c. f. its title), and finding a focus in itself is a major part of this work. Especially from the last part of this exercise, centered around works directly related to urban and geovisualization, potential avenues for further investigation can be seen to crystallize. Criteria for narrowing down paths to continue work on are the *frequency*, persistence, or importance of problems (how often and urgently is a problem being brought up?), *intersectionality* (can this problem be tackled with one or more advances from elsewhere?), and *feasibility* (can it reasonably be expected to be approached within a thesis?).

The most frequently stated problem is also the most general one: being overwhelmed with (spatial) data that is growing in size, detail, and heterogeneity. *Size* can of course mean the sheer data volume, as measured in bits, but more relevant to this thesis is the meaning that is closer to *detail*: spatial data that is getting finer-grained in its resolution, while spanning the same or larger regions — in other words, a problem of *scale*. *Heterogeneity* contributes to that, in that for each point there is now also an increasing number of different measurements, which can not always be trivially combined into a single visualization. In other words: making sense of geospatial urban data is a *multiscale* and *multilayer* problem. Since this thesis is not directed toward geographic analysis, visual sensemaking of this data is the aim.

The work discussed in this chapter contains approaches that are suitable here: geovisualization and visual analytics already use or propose multiple coordinated views, which can show different layers of spatial data side-by-side; or different areas or scales at the same time, similar to a minimap. Nonuniform displays can perform a similar function, but are instead merging multiple layers or scales into a single view that is interactively altered. Recent work also increasingly brings up the potentials of immersive technologies, and how they can enhance visual analytics and exploration for a broad range of uses. A feasible question for this thesis, aiming to tackle a common problem by building on interdisciplinary methods thus seems to be: *how can immersive technologies be employed to improve the visualization of multilayered urban data, and make multiple scales easier to navigate?*

Part II

Hypotheses and Methods

3 Immersive Environments: Engaging Space for Spatial Data

In order to get an understanding of what *immersive visualization of urban data* is, the previous chapter went on a journey through published work on the individual parts that make up this broad subject, and through related fields that could point to where it can go with the possibilities on offer. This initial research already opens a possible hypothesis for this thesis to test: *immersive environments have the potential to improve visualization of multilayered urban data at multiple scales.*

In addition to purely literary research, another method offered itself to either strengthen, or perhaps expand this hypothesis: asking urbanists themselves what their process is, how they work with (three-dimensional) urban data, and what, if anything, they would like to see improved in relation to its visualization.

3.1 Engaging Urbanists?

In an attempt to find out more about the urban design process beyond what a cursory glance at some literature can reveal, and to also see what its current and local state is (as of 2017 in Nantes, France), we decided to hand the task of interviewing a diverse cast of local actors in urbanism to a master student interning at our research team. Amandine Dugrain devised a semi-structured interviewing scheme that aimed at clarifying these questions:

- What are the phases in a typical urban development process?
- Who are the actors involved in it?
- How much are these actors involved at each step?
- How, and how much is 3D visualization or data used?

From her work of analyzing each interview she conducted with around 20 participants, the chart shown in Figure 3.1 resulted as an answer to the first three questions. This chart offers a look at the length, complexity, and diversity involved in urban development, showing a large number of different roles interacting with each other at many different

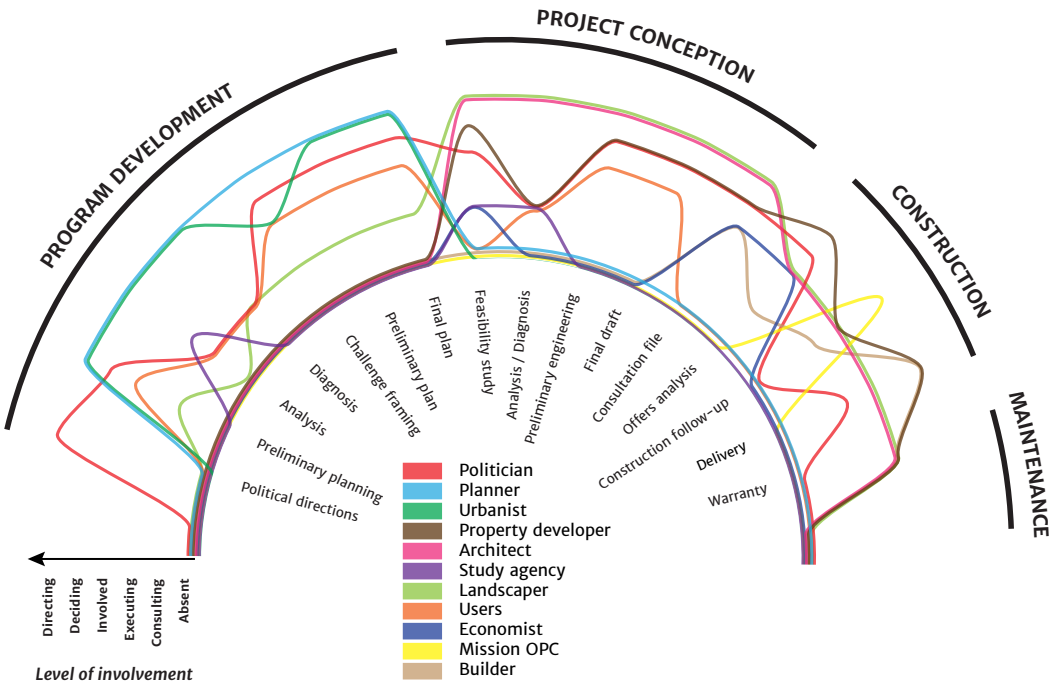


Fig. 3.1: The different steps and phases in a typical urban development process, and the level to which different actors are involved in them.

steps. The frequent concurrent involvement of different roles at many points implies a necessity for communicating analyses, plans, and progress updates between diverse actors. From this it can be taken that tools and methods used for such communication can not be too niche or require too much specialized expertise in order to be useful to as many participants in the urban design process as possible.

To the question about what kind of 3D urban data was employed and how, the answers painted a less clear picture. At least from the experience of those professionals who were interviewed, their use of 3D visualization[1] appeared to be quite sparse. Its main purpose was reported to be to communicate plans and proposals in the form of rendered 3D models, with the aim of providing more engaging images. Contrary to the plentiful possibilities from much of the literature cited above, work within the local urban planning offices mostly happens with the aid of printouts of various 2D maps laid flat on large tables or hung on walls, without using advanced visualization systems. This could speak to a rather simpler nature of the projects our sample of professionals were dealing with, but when asked directly about this apparent absence of 3D data, its unavailability was cited as a main reason, along with inadequate ease-of-use of the software required for it. That is not to say that an aversion to 3D itself exists: the interviewees were in fact

[1] '3D data," "3D visualization," just "3D," etc. are used almost interchangeably here. This is in part a reflection of the vagueness with which the interviews were conducted, but also of how little importance anything "3D" was given by most interviewees, based on its lack of use. In this context, these terms can mean any of the following: 3D renders of architectural or urban designs; 3D urban data, such as building shapes and underground utilities; and interactive systems that visualize such data in a 3D environment.

very much eager to adopt 3D data and visualizations more widely, if it became more accessible.

While these interviews did result in an interesting overview over the actors in urbanism and their involvement in urban development projects, their other goal was not quite accomplished: there were no concrete wishes or specific problems the interviewees expressed in relation to visualizing 3D data for their purposes. On the other hand, the takeaway from this, in combination with the expressed eagerness to use more 3D if became "easier to use" is a kind of *carte blanche* for this thesis from that standpoint: the creation of anything appealing in terms of 3D visualization might be welcomed by the urbanist community, and could foster wider adoption of 3D at a number of stages in the urban design process — the benefits of which are clear from the overviews in the previous part. Also clear from the previous part is how immersive technologies can be more *engaging* than traditional displays when used for comparable tasks: examples of this exist in fields as diverse as software development [Mer+17] and climate change education [Que+18], and interactive virtual reality has been shown to be highly engaging and supportive of exploratory and analytical activities [Lu+18].

For the purpose of creating something that could be beneficial to the diverse cast of actors in urbanism, immersive environments thus appear very worthwhile to investigate: their engaging nature could even make existing 3D systems attractive enough to widen their adoption. Of course, beyond making simple 3D maps more engaging to use, immersive environments offer additional benefits over conventional display technologies, and open possibilities that could not be realized otherwise, such as arbitrarily large workspaces.

3.2 Virtual Space

Large, high-resolution interactive displays have seen increased interest in visual analytics research since LCD panels that surpassed CRTs started becoming commercially available during the 2000s. While ultra-high-definition devices such as 8K television displays and monitors have only been released quite recently in the mid 2010s, tiling individual, smaller panels allowed to create composite displays that could exceed the size and resolution of modern singular ones many years earlier.

Such arrangements were found to "improve performance time for basic visualization tasks in finely detailed data" significantly over regular desktop displays in a study by Ball and North [BN05]; comparing grids of 4 and 9 monitors and a single one. Given the tasks of finding and comparing small targets in high-resolution images, participants would have to pan and zoom on smaller displays to explore the full image space; only the largest configuration could show all at once. Their participants by far preferred

physically navigating around the larger monitors over interacting with the smaller ones virtually, with one explanation being that they preferred not to "lose context:" panning and zooming would move parts of the image beyond the display frames. The experimenters also observed the use of larger displays to be less stressful, and to result in a "better sense of confidence" over the smaller ones. The same authors have later shown how the physical navigation afforded by larger displays, i. e. moving eyes, head, or body has also been shown to be advantageous over virtual navigation [BNB07], being more efficient and strongly preferred by users for pattern finding, search, and navigation, and insight-gathering tasks for spatial data.

Andrews et. al. [AEN10] performed more complicated studies aimed at comparing a large "analyst's workstation" (a 4 × 2 grid of 30-inch LCD panels with a total resolution of 32 megapixels) with a then-typical 17-inch monitor to assess how display size can affect the "sensemaking process" in analysis and synthesis tasks, with both students and professional analysts. The larger display was shown to support sensemaking by "becoming part of the distributed cognitive process, providing both external memory and a semantic layer." Even participants who initially dismissed the larger space as "way over the top" and doubting the use of more than one monitor were observed to eventually leverage the full display space to their advantage and expressing enthusiasm about it. Notably, their large display setup is intended for a seated position — physical navigation is limited to glancing around and turning the head, yet still the positive effects of spatial memory [Kir95] and not having to switch context were possible. The authors later outlined the "issues, challenges, and opportunities" with large, high-resolution displays for information visualization [And+11]. A very pertinent quote from that work:

> Large, high-resolution displays are fundamentally different from conventional displays. Rather than being small portals in which the user must fit their work, they are human-scale environments that are defined more by the abilities and limitations of the user than by the technology. Designing for these displays is thus not simply a matter of scaling up existing visualizations or displaying more data; instead, designers must adopt a more human-centric perspective.

An example of using the potential of large, high-resolution displays in this way is proposed by Isenberg et. al. [Ise+13]: a "hybrid-image" visualization that blends "two different visual representations into a single static view, such that each representation can be perceived at a different viewing distance." Standing at some distance, users have an overview and can see much finer details by moving closer, again without losing context. Using high-pass and low-pass filtering of the near and far images, respectively, and alpha-blending them results in the different levels of the visualizations not interfering with each other visually. The authors demonstrate this approach with dual-scale treemaps,

network diagrams, scatterplots, and hybrid visualizations that present different data representations at each scale. They conclude that this approach represents a "solution for providing viewing distance dependent overview+detail views in a single image," and note how the hybrid-image technique makes designing such visualizations easier by making it possible to combine existing visualizations "designed for different viewing distances into a single image."

Large, high-resolution displays continue to be explored for applications in visual analysis, where their benefits on knowledge discovery and insight acquisition keep getting confirmed [Red+15]. They are now considered as one of the archetypal systems for immersive analytics [Büs+18], which also include CAVE-style environments [CSD93], which employ projection or more recently large displays in a room where the user's head is tracked to provide a virtual reality environment with a large field of view, and of course head-mounted virtual and augmented reality devices.

In these immersive environments, basically everything around the viewer is a kind of "display," by definition. Virtual reality envelops a user in a fully artificial environment which can be filled with any 3D models, including virtual 2D screens that emulate traditional monitors, but without any restrictions on size and layout. With the increasing resolutions and field-of-view that commercial VR headsets offer (115 degrees, and "over 70" pixels per degree in the case of the central part of the latest devices [Tec]), these virtual screens can offer all benefits of the large, high-resolution displays discussed above, at a potentially lower price and without the physical requirements of actual display units. Restrictions on positioning and interacting with these objects are imposed only by the physical room within the user is situated (walking into walls or furniture) and, if using a tethered VR HMD, the length of headset's cable. Instead of requiring direct touchscreen input, VR devices usually already incorporate flexible controllers that are tracked in 6 spatial axes, potentially offering more ways to interact with visualizations than large physical screens. Augmented reality can be seen as being more limited in that it should take into account the physical environment, so as not to intersect both "worlds," but on the other hand it can integrate reality in potentially useful ways, such as augmenting real objects with added information, and extending existing and familiar workspaces with virtual monitors, as proposed by Ens and Irani [EI17].

Ens et. al. previously explored design spaces for placing multiple virtual windows, or "Ethereal Planes" in virtual environments [EFI14; EHI14]. Through experimentation with their "Personal Cockpit" design they arrived at a curved multi-window layout that was optimized for fast task switching between windows — a design that would be highly impractical using physical monitors. While many of the works cited above reference the advantage of physical over virtual navigation, the same movement can lead to discomfort, especially the movement of the neck that is required to view curved displays arranged around oneself. This is however yet again another area where virtual workspaces of-

fer solutions that would be impossible with physical displays: McGill et. al. propose an automatic counter-rotation of the displays, mapped to the user's head movement [Mcg+20].

These works strengthen the case that immersive environments would be an excellent domain to further explore visualization of urban data in. Not only are they engaging [Ito+16; Dwy+18] and in high demand (as seen from the continued commercial interest and adoption), but can offer the same benefits that large, high-resolution displays have been shown to possess for visualization tasks, and even exceed these with layouts and interaction methods that would be hard to match in the physical world, such as effective multi-monitor setups. Of course, such virtual information displays can also add features that are impossible otherwise, such as expanding data into the third dimension [Liu+20] and linking displays in 3D [RFD20], or even add complex spatiotemporal visualizations in a space-time cube with tangible controls [FSN20]. This last example reference already targets geovisualization (specifically, trajectory data exploration), using one large virtual display in a reproduction of an "analyst's desk," and another recent one explores 3D multiview map layouts for the same above reasons: immersive spaces "are not limited by the real-estate barrier of physical screens and support sensemaking through spatial interaction" [Sat+20].

The two challenges with urban data identified in section 2.4 thus seem to be perfectly approachable with these strengths of immersive environments. Large, high-resolution displays have been shown to be able to address issues with *multi-scale* visualizations, and virtual versions of them are potentially more flexible and effective. The advances shown for immersive multi-monitor setups can be used to address the *multilayer* challenge, if each layer is shown separately in its own (virtual) window. The first hypothesis from the beginning of this chapter could thus be stated more precisely as follows: *the virtual space and interactivity afforded by immersive environments offers solutions to visualizing multilayered urban data at multiple scales, while being highly engaging.*

4 Data Layer Separation: Immersive MCVs

The just mentioned hypothesis contains a plural in its formulation: "immersive environments offer *solutions* to multilayered urban data at multiple scales." To address these two challenges, it may be best to approach them separately. *Separation* is also what forms the basis for the next hypothesis: *the virtual space afforded by immersive environments allows for novel arrangements of multiple and coordinated views to create effective visualizations of multiple layers of spatial urban data.*

4.1 Genesis: Taking the Data Layer Concept Literally

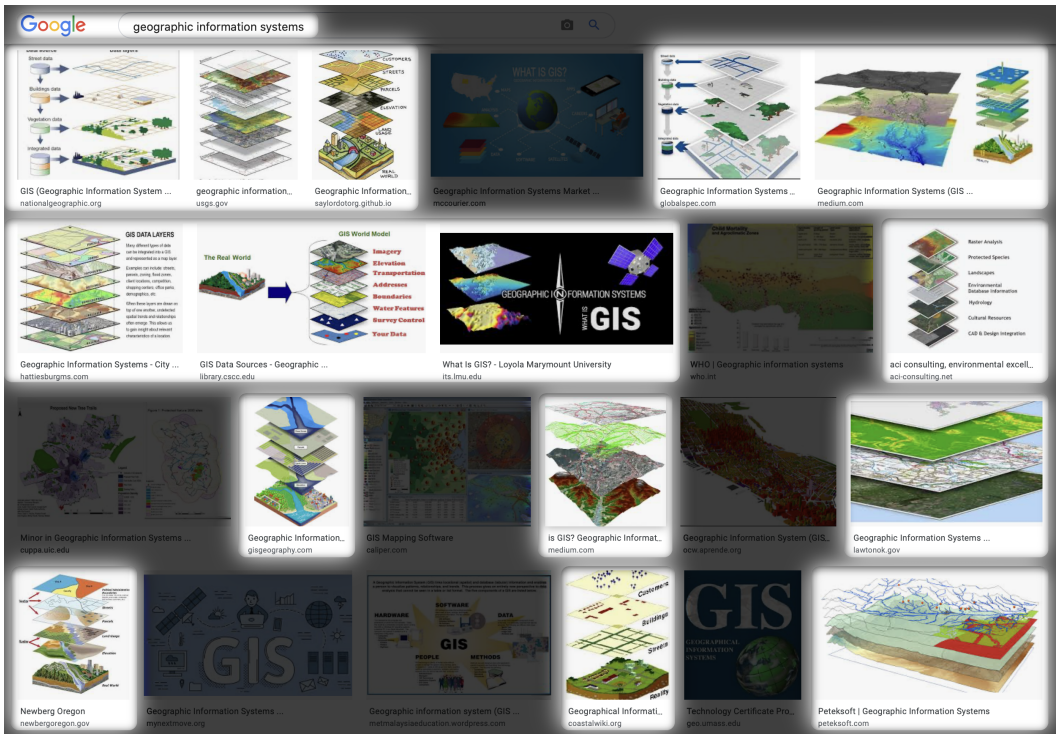


Fig. 4.1: A Google image search for the query "geographic information systems;" highlighting the majority of results which show multiple data layers illustrated as a vertical stack.

When doing an image search for "geographic information systems" (Figure 4.1), the majority of at least the first dozens of results are illustrations of how such systems accom-

modate multiple layers — one of the fundamental concepts of spatial data organization [LT92]. Such images are ideal to convey the concept: square slices of a landscape, closely stacked and often overlapping signify their spatial coherence even if the layers may be visually very different, and because of their (usually axonometrically) projected tilt, even illustrations of a relatively large number of layers do not nearly take up as much space as they would if juxtaposed side-by-side, without projection. This principle also applies not only to thematic layers, but also to temporal slices showing the same place and data type at different, and to literal slices of vertical structures such as buildings.

The advantages of separating layers over attempting to merge all data into one complex map are clear, especially in 3D: buildings and terrain lead to occlusion, too many items lead to crowding, and the number of available visual channels is limited for how much information can be legibly overlaid. Clever cartography can alleviate these issues to some degree and continues to be an area of intense research, but the kind of effort required to execute this effectively is itself another downside of single maps over a layered approach. Obviously, single maps are maximally compact in their overall (display) space requirement, and in not distributing spatially coherent items across separate views, but here the stacked layout is offering a compromise.

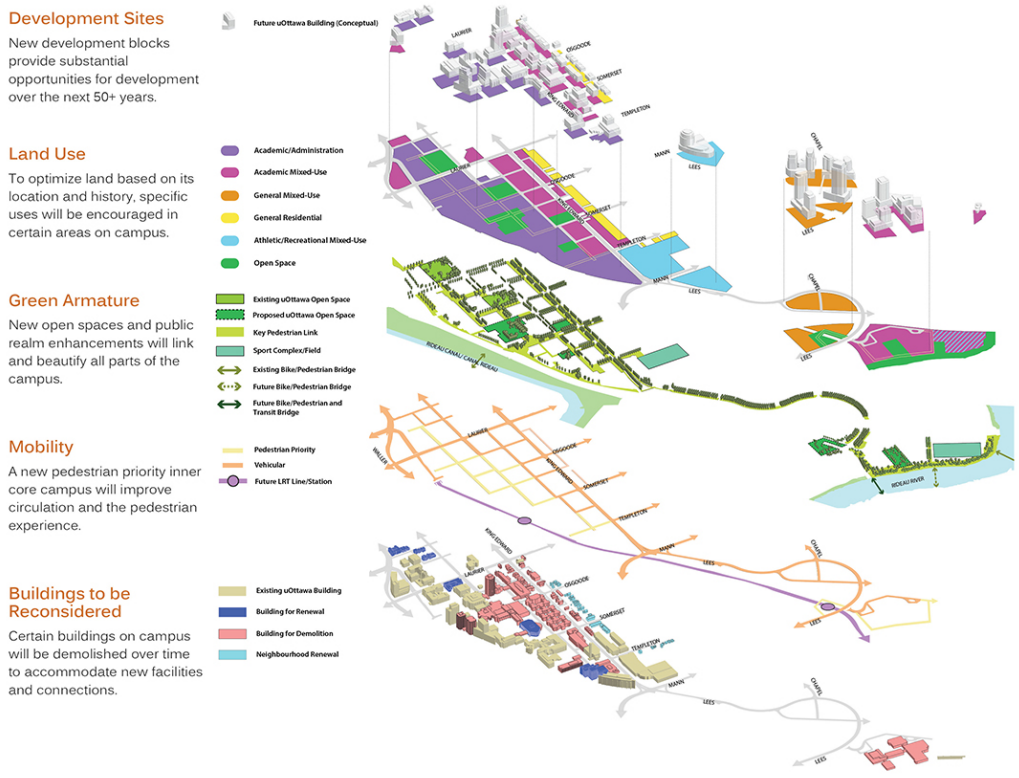


Fig. 4.2: A vertical stack of layers of the University of Ottawa's campus master plan.

Given these benefits of keeping layers separate, and the at-first-glance effectiveness of using vertical stacks, this way of visualizing multiple layers is rarely used outside of

roughly illustrating the concept, or at most listing which particular data layers exist for a certain application. A notable exception where this "stacked layers" approach is applied to actually communicate detailed information within each layer is urban development master plans, such as that of University of Ottawa's, shown in Figure 4.2. This example shows the utility of the method: each layer only contains as much information as is clearly legible, yet together they display a coherent picture of all data. For example, Layers like "Green Armature" and "Mobility" would overlap, and the 3D buildings would obscure information below or behind them. It also shows the limitations. This is a static image, hand-crafted to optimize for the particular angle that is used here. The layers could be as close together without overlapping only thanks to the elongated shape of the visualized area, otherwise a much greater tilt would be required. The question is: could this concept of a stack of geospatial data layers be useful as an interactive, dynamic visualization?

4.2 Stacked Layers: Related Work

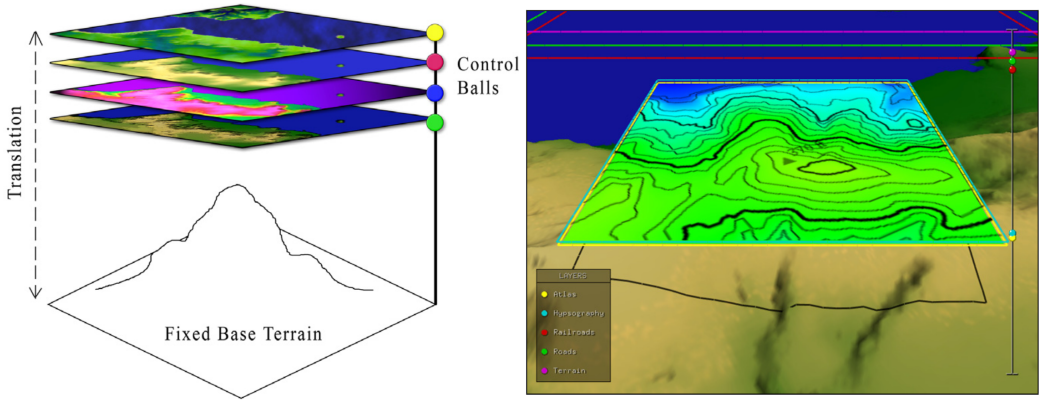


Fig. 4.3: The concept and implementation of Brooks and Whalley's 3D GIS layering system [BW05; BW08].

Brooks and Whalley have introduced a 3D GIS visualization system that appears to implement this layer stack concept in an interactive fashion [BW05]. In their proposal, the 3D base terrain of an area is shown, and above it float square segments of the region, displaying different data layers. These layers can be freely placed vertically, with sparse layers such as road or hypsography layers being able to be merged with textured layers. As layers are raised from the terrain, they are gradually flattened from their original digital elevation model. While the description of their system presents basically an interactive and useful form of the layer stack concept used for many of the illustrations seen in Figure 4.1, all examples (screenshots) of their work only show a single textured layer floating above a larger terrain, all resembling variations of what is seen in Figure 4.3.

In the latest work featuring this 3D hybrid GIS [BW08], the authors mentioned their intention to "undertake a usability study," however literature search did not yield any continuation of this project, nor any similar research on multilayer separation that

resembles the stack illustration for urban data, or GIS in general, though there exists some related work that is worth mentioning.

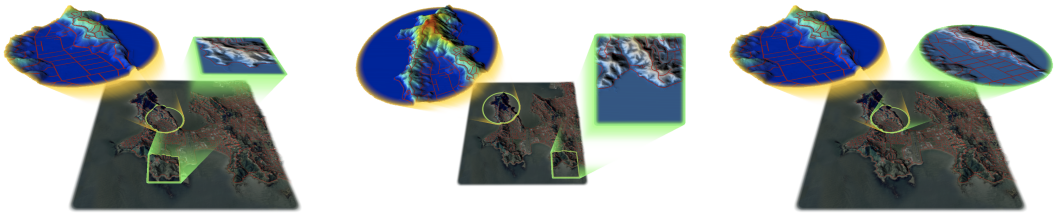


Fig. 4.4: Examples of Trapp and Döllner's rendering technique from [TD19].

Trapp and Döllner present a rendering technique [TD19] that most closely resembles Brooks and Whalley's multilayer hybrid, in that it shows a 3D terrain base layer, and projects certain segments of that map above it. As demonstrated in Figure 4.4, their system can project multiple such regions, and the projected "3D close-ups" can feature different "data mappings," including multiple separate projections of the same region using different mappings. As their system is primarily intended for *detail+overview* visualization, the projected segments are magnified and moved away from obscuring the base terrain to enhance visibility of both; they are also not stacked. Neither are multiple data layers used; only different rendering schemes (mappings) of the base layer that emphasize certain qualities, such as elevation. Their work and its evaluation is focused on the technical aspects of their algorithm and its performance on GPUs — a user study has also not been attempted here.



Fig. 4.5: View of the vertical separation of streets and points of interest from the urban base map, as investigated in [ED15].

Closer to the layer stack concept is the work of Edler and Dickmann [ED15], where parts of a map (streets, points of interest) are separated and vertically elevated above it. This applies to the whole map, as in the illustrations from Figure 4.1, but the elevated layers are mostly transparent, thanks to their sparse nature. Since occlusion is thus minimal, the layers are kept relatively close to each other to preserve coherence between them.

The goal with this technique is actually to just *accentuate* the streets in order to improve the speed of map reading. This accentuation occurs by their system being shown on an autostereoscopic display: seeing the elevation in three dimensions makes it clear as such, instead of possibly mistaking it for just shifting the street layer horizontally, which this technique could be mistaken for by just looking at a two-dimensional representation as in Figure 4.5. Their study, in which participants were asked to count certain symbols as quickly and accurately as possible, showed significant improvement in the speed with which the tasks were completed, and a slight improvement over the accuracy, as compared to doing the same task with the layer elevation disabled. This confirmed the authors' previous work [Edl+14] on "spreading map information over different depth layers" in order to improve map-reading efficiency, although this improvement is being attributed to the spatially accentuated streets now better subdividing the map into different "spatial chunks." This kind of division has previously been shown to influence cognitive information processing in spatial layouts [MRM84; Bes+13; Hur05].

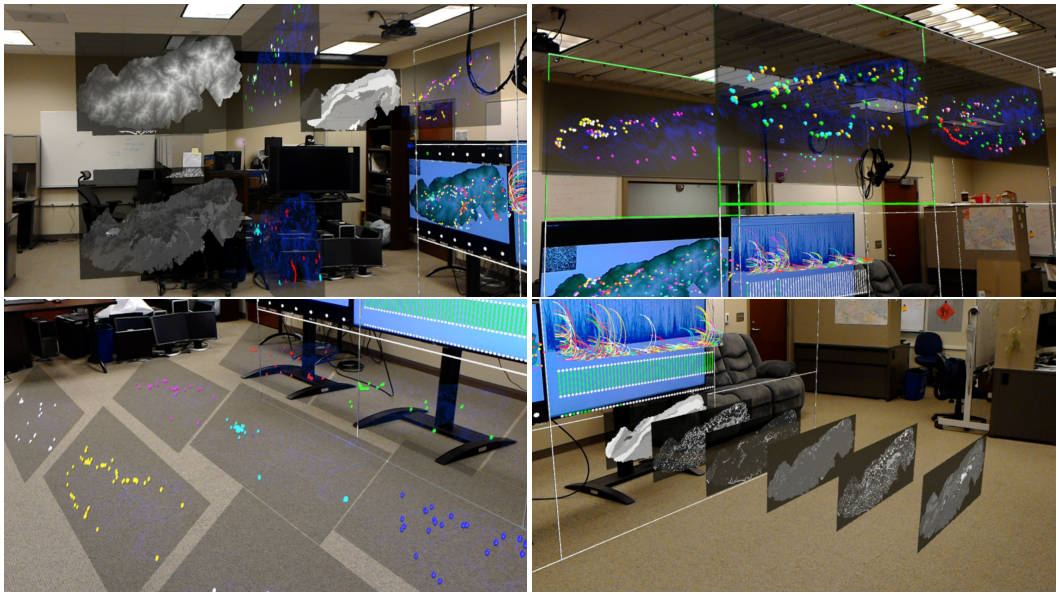


Fig. 4.6: Examples of the Multiple Coordinated Spaces from [Mah+18].

Finally, the map layer stack can be seen implemented as a part of Mahmood et. al.'s *Multiple Coordinated Spaces* system [Mah+18], however stacked horizontally, as shown in the lower right corner of Figure 4.6. In fact, their system is built on the idea of being able to position visualizations in many different spatial configurations around any room, with an emphasis on integrating them with existing flat surfaces and distributing them in 3D grids, as seen in Figure 4.6's examples. The stacking of "holographic maps" is not the main point of this work, and has been treated only tangentially there. Their focus was rather utilizing the *distributed cognition* theory [ZN94], and how it can be applied to information visualization [LS10] with large displays of multiple coordinated views in a flexible, immersive environment. They make similar arguments about how immersive technologies create new opportunities for visual analytics and sensemaking through their

"new interaction methods mixing physical and virtual operations" and "large rendering space with unlimited pixels," which allows "cognitive offloading" for complex tasks, also citing work by Andrews et. al. on large displays [AN12].

As seen from Figure 4.6, their MCS system does much more than just show different layers of a map: it visualizes the available data in different representations and links them across the views. In their case studies, they used data on biodiversity from the Great Smokey Mountain National Park, which includes the habitat locations and geo-distributions of 869 plant and animal species, and their interdependencies and correlations with 59 geo-distributed environmental variables. These are visualized as heat maps, scatter plots, parallel coordinate and edge bundle graphs, and all are connected: filtering or selecting data in one visualization updates the others.

While not conducting a quantitative experiment, the authors do describe in detail how their system can be used in an effective way for analysis, exploration, and discovery based on their case studies, mentioning multiple times how this achieved in part by spreading out or repositioning their "holographic layers" in space. As mentioned before, the main focus is however on taking advantage of physical rooms and integrating 2D visualizations in 3D space, and coordinating the different visualizations. The *distributed cognition* approach seems to take precedence here over immediate, concurrent visibility of all data — the system actually encourages users to walk around the room from display to display during investigation. Such "intelligent use of space" [Kir95] is well founded in research on cognition and interaction: "physical navigation engages more embodied resources, providing a greater opportunity for users to couple with the space, providing more meaning to locations and encouraging a more cohesive view of the entire workspace" [Dou04].

4.3 Map Composition and Comparison

While the exploration of using physical space to such a degree that it requires traversing the room is intriguing, the original map layer stack concept from its illustrations (Figure 4.1) is attractive in its compactness and immediacy. These qualities seem to make it an ideal candidate to investigate for its effectiveness as a *map composition* technique for *map comparison*.

In cartography, a main desire as well as a major challenge has long been to mix heterogeneous geographic data [HCM13]. Different data layers can be combined into one map by utilizing various *visual multiplexing* techniques [Che+14] that e. g. engage different visual channels, and advances have been made that can take semantic attributes of data layers into account when partially compositing them [LAP17]. A way to sidestep this

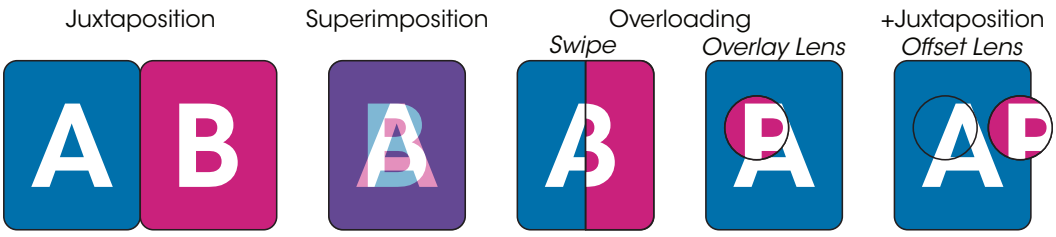


Fig. 4.7: Layers "A" and "B" composited using different strategies. Not shown: *blitting*, or *temporal multiplexing*; where the layers are interactively switched in place.

often laborious process with rarely an obvious solutions is to use more simple *visual composition operators* [JE12], mainly *juxtaposition*, *superimposition*, or some forms of *overloading*; illustrated in Figure 4.7.

Lobo, Pietriga, and Appert performed a detailed evaluation of five map comparison techniques for two layers based on these operators [LPA15], with systems similar to the illustrations in Figure 4.7:

- **Juxtaposition:** showing both layers next to each other, displaying the same geographical region in both views,
- **Translucent Overlay:** overlaying or superimposing both layers with an adjustable opacity of the upper layer — immediately switching between 0 and 100 percent opacity is equivalent to *blitting*,
- **Swipe:** both layers overlaid, with a divider that can be swiped horizontally to reveal the lower layer,
- **Blending Lens:** a movable lens through the overlaid upper layer, with variable opacity, and
- **Offset Lens:** a movable lens that shows the underlying layer in the selected region of interest, but offset from its original position — a localized juxtaposition.

The authors discussed the advantages and disadvantages of each system in terms of visual interference, divided attention [Har+95], and scanning strategy (motor-driven vs. vision-driven), before conducting their user study:

- **Interference:** Juxtaposition and Swipe have zero visual interference, as they present both layers fully intact without blending them together, whereas the Translucent overlay can suffer significantly from it, depending on the map's contents. The two lens techniques only interfere inasmuch as a portion of the map will be blocked by the lens.
- **Attention:** Juxtaposition produces the most divided attention, as users have to repeatedly inspect each viewport. Reducing distance between layers lessens this division; it is zero in the full and lens overlay.
- **Strategy:** Swiping and the lens methods require the most motor-driven approaches, as revealing the other layer needs interaction. The overlay's main motor action is

changing the opacity, and can be further reduced by making it a binary choice; it thus becomes the mostly vision-driven. Juxtaposition similarly is mostly vision-driven, but its reduced view ports (if using the same window size for all methods) requires more panning interactions.

In a study where users were asked to compare topographic maps and their corresponding orthoimages for differences (one of the *objective primitives* from Roth's taxonomy of interaction primitives for interactive cartography and geovisualization [Rot13]), hypotheses based on the above characteristics were tested. The five systems were evaluated for their difficulty and completion times in finding multiple kinds of differences in various geographical features. In their findings, juxtaposition performed only well for comparing facilities as opposed to roads, but not significantly better than the blending lens or the overlay. Otherwise, the translucent overlay has been found to be the best overall technique, and is only bested by the blending lens in tasks entailing an identification of modified or added entities — both techniques complement each other. The swipe and offset lens techniques performed the worst, apparently due to their tight coupling of motor actions and visual comparison: an offset lens still divides attention while requiring significant movement of the lens, and swiping forces users to either constrain their scanning or to execute many micro-swipes back and forth. This calls into question the recent prevalence of swipe interfaces in web applications and the floating lenses [TD19] shown in Figure 4.4, at least for applications comparing two maps.

The authors note how these five techniques "have never been compared empirically on well-defined map comparison tasks," and related work that they reference only evaluated one compositing technique [LM14b; Bon+09; RB11], just one aspect of a technique [Har+95], or was geared towards showing different representations of one map and not of multiple layers [PW06; JGE12; Pin+12]. These kinds of evaluations in geovisualization — comparisons of techniques to display multilayered maps — remain sparse, with this work by Lobo, Pietriga and Appert [LPA15] still being the most comprehensive example of its kind. Even so, it (like most other works) only covers techniques meant to compare two layers of a map, using a traditional desktop environment. Would these findings be applicable to a higher number of layers, and to immersive environments?

The same authors later introduced a layer compositing system [LAP17] that incorporates their previous findings, and provides effective means to dynamically composite multiple map layers employing different techniques. This approach is aimed at GIS experts and potentially at "lay users who want to design custom maps," which can be seen in the complexity of a screenshot of their proof-of-concept implementation (Figure 4.8). A comprehensive system like this can hardly be compared to individual techniques, as it is itself a collection of such. It does however highlight one of the main problems in selectively adding layers beyond just two: the need to individually manage each layer and its presentation and placement on the base map. The same can be observed on even

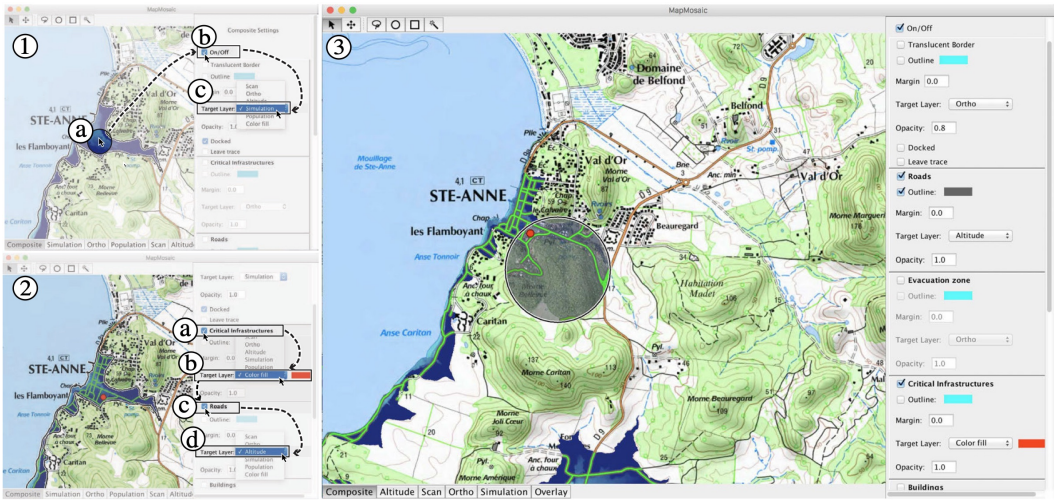


Fig. 4.8: Example workflow of creating a composite map from different layers using Lobo et al.'s MapMosaic [LAP17]

the much simpler approach of Trapp and Döllner's interactive detail+overview lenses [TD19] (Figure 4.4), if those lenses were to be used as offset lenses similarly to Lobo et al.'s previous work [LPA15]. Each "focal point" of each lens needs to be moved, as well as the projected images themselves to prevent overlaps whenever a different area is to be revealed. The number of operations thus necessary increases with a multiple of each new layer added, unless lens placement is automated, which itself could cause confusion, unpredictability, and frustration.

Swiping, already evaluated as less favorable by Lobo et al.for interacting with just two layers would quickly become even less usable with a higher number, but so would the overlay for any number of non-sparse, or textured layers, as e. g. blending their colors could make their content unreadable. If different opacity values needed to be managed for each layer, the number of possible configurations would also increase geometrically with each new layer, burdening the user with finding optimal settings even more.

These considerations leave only juxtaposition as a viable technique for *visualizing* maps with more than the smallest number of multiple layers in an uncomplicated way. The emphasis here being on using maps in their primary roles, as per MacEachren's *Roles of Maps*: exploration, confirmation, synthesis, and presentation [Mac11], and not on building e. g. authoring tools to create new (single-layered) maps from multiple layers. This includes juxtaposition in the temporal sense, as discussed by Gleicher et al.: "alternat[ing] the display of two aligned objects, such that the differences 'blink' "[Gle+11]. Lobo et al. refer to this as temporal multiplexing, or blitting, and consider it as a version of binary superimposition. Blitting may in fact already be the most common composition or comparison technique, as it is the only method available in most lay-person oriented mapping software (e. g. switching between road networks and aerial imagery on a smartphone app), and a most commonly used one in GIS software, even if just for

"flipping through" the available layers. In this manner, doing the same with printed pages in an atlas could also be considered a form of blitting.

Juxtaposition, in this sense that it presents mostly equal views of different layers (in a similar view, with a equal extents) also opens up the opportunity of employing techniques from the area of *multiple and coordinated views* (MCVs) [Rob07], at least in some of its simpler forms. Examples were already employed in the map comparison techniques implemented by Lobo et al.: a cursor is mirrored in their juxtaposition view over both maps, and every composition technique pans and zooms both maps simultaneously — the user only has to interact with one, and does not need to realign different layers. This kind of linking can of course be applied to an arbitrary number of juxtaposed map layers, which would necessarily get smaller to all fit in one window, and increase the distance between them as their number grows, which is one of the major downsides of juxtaposition as remarked by Lobo et al., contributing to the problem of divided attention. This is however where a functional, vertical stack of layers could be a solution.

4.4 Stacking Layers Vertically as MCVs — in VR

Comparison visualizations more generally and their considerations are discussed by Gleicher [Gle18]. In an earlier paper [Gle+11], he posited "three basic designs for visual comparisons:" juxtaposition, superposition, and explicit encoding (a direct visualization of relationships). Other methods, such as the overloading and nesting explored by Javed and Elmquist [JE12] and evaluated by Lobo et al. [LPA15] can be seen as combination of these basic ones. There are a number of takeaways from Gleicher's considerations that are confirming the idea of a vertical stack, and guide certain aspects of it:

- **Many objects:** "Superposition becomes cluttered with many objects in the same space. Juxtaposition with many items separates them, hindering some kinds of comparison. Either approach requires some strategy for managing the complexity, although juxtaposition may naturally support scanning."
- **Layouts:** "The three-axis space [number and size of items, and complexity of relationships] focuses on a particular aspect of visualization design: the layout of the presentation. Careful design choices are required to make any of the three design strategies work. [...] For juxtaposition designs, having similar layouts can reduce the burden on the viewer."
- **Interaction:** "Interaction is often a mechanism for addressing scalability concerns in visualization. It applies across the space of comparative designs. Interaction can be used as a mechanism for addressing issues with a design strategy, for example, using brushing and linking to help establish connections between juxtaposed views."

For "many objects," i. e. multiple maps, juxtaposition is as discussed above necessary to avoid clutter, occlusion, and overly complicated usage. It however introduces a necessary distance between them, that Lobo et al. say causes divided attention. To better support the strategy of *sequential scanning* [1] and to minimize juxtaposition's downsides, an appropriate *layout* needs to be chosen — ideally one that supports "good ordering," which is "critical to the success of scanning strategies." Ordering can "help address complexity by placing related items together" and enable users to finish comparisons sooner [Gle18]. A typical arrangement would be a grid, the dimensions of which would depend on those of the individual views, their number, and the size of the display.

In a grid of map layers, all showing the same area at the same size, the distance between the same positions on different layers is at least as big as the dimensions of the layer views themselves, plus a potential separating frame between them. Arranging them in in a vertical stack can bring this distance down considerably, but also introduces other challenges and opportunities along with this, including opening the possibility of directly showing terrain and other three-dimensional features on the layers. Vishwanath, Girshick, and Banks [VGB05] have shown how the perception of images *with binocular viewing* is "nearly invariant across a wide range of viewing angles" such as those a stack would introduce, so trading this additional tilt for a significant decrease in distance seems worth investigating. Furthermore, a tilted view is by far the most common way maps are have been and still are presented: as printouts laid flat on a desk. This horizontal layout continues to be used with more recent technology, such as interactive tabletop displays for geospatial networks [NDV12], and immersive space-time cube visualizations that use a map on a virtual desk as its base [FSN20].

In Figure 4.9, an example of a 2×2 grid of four layers, and two stack arrangements, is seen sideways. Only layouts in 3D space are considered here, not isometric projections such as can be seen in Figure 4.2, as those would not work well in immersive environments. While the two stacks (Figure 4.9 b and c) would need to be displayed in an immersive environment or be projected on a screen between the viewer and the stack, the grid (a) as shown here could be a monitor screen itself, as is already the typical use of it.

Figure 4.9 (b) shows a "flat stack," similar to the prevalent illustrations of map layers (Figure 4.1) and as implemented by Brooks and Whalley [BW08] (Figure 4.3) and Mahmood et al. [Mah+18] (Figure 4.6). Similar to the grid in Figure 4.9 (a), the regular arrangement and spacing of the flat stack preserves relative distances across layers, however much closer together than in the grid thanks to its tilt. Also as in the grid, the actual appearance to the viewer of each layer will be different: the same point on each layer will be viewed at a different angle, from a different distance (the illustration uses different colors to highlight dissimilar lengths). This kind of stack also introduces possible occlusion of

[1] The other two general strategies for comparisons from Gleicher's considerations being *subset selection* and *summarization*, which are not directly applicable to displaying multiple layers of maps.

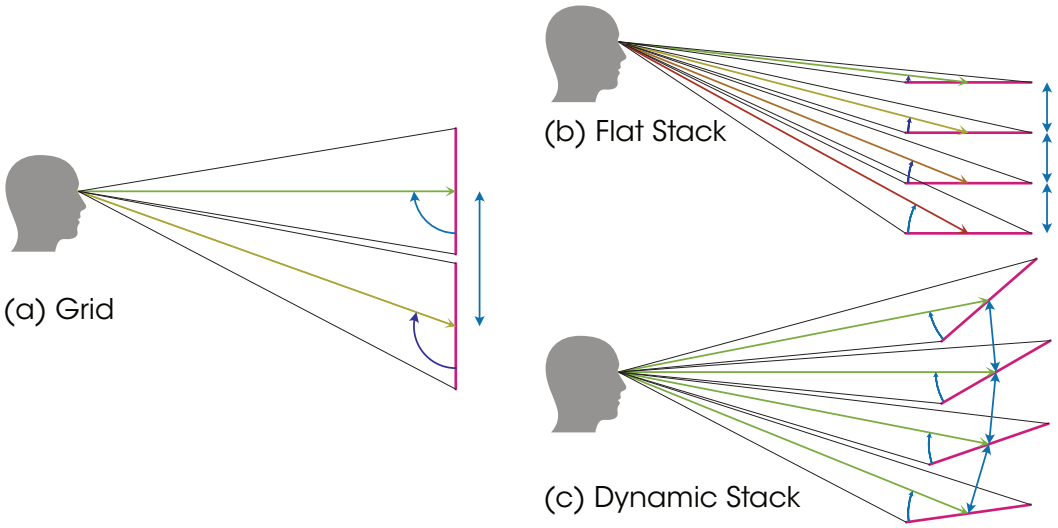


Fig. 4.9: Side views of 4 map layers arranged in a 2×2 grid (a), stacked vertically in a flat arrangement (b), and in a dynamic stack that tracks the viewer's position (c).

one layer by another (the far end of the bottom one is obscured by the one above), and requires all layers to be situated below the viewer to be visible at all — the top layer in Figure 4.9 (b), close to eye level is already at a very sharp angle and would be barely legible, especially if it included 3D features which would themselves strongly occlude information behind them.

All these downsides can be solved with a "dynamic stack" that tracks the viewer's position and adjusts the layers' position and orientation, or a static approximation of such if the viewer can be assumed to not move around too much. [2] Such a stack (Figure 4.9 (c)) can keep each layer at exactly the same distance from the viewer, at the same viewing angles, and the same distance from each other, so as to avoid occlusions. This way, each layer appears at the same size with the same perspective distortion — this visual similarity can "reduce the burden on the viewer," according to Gleicher [Gle18]. With each layer appearing the same way to the viewer, the dynamic stack is also very close to the axonometric views that illustrations such as Figure 4.2 provide, which themselves would appear distorted in a stereoscopic environment.

In this dynamic system, the distance between the layers could be quickly adjusted by users by either changing the visual gap between them, or setting different viewing angles — the other parameters would be recalculated according to the constrains above to preserve this visual similarity between layers, whereas changing the same on a flat stack could result again in occlusion and loss of visibility, and would require additional care from a user, potentially needing to adjust each layer individually.

[2] In immersive environments the exact location of the viewer's head is by necessity always known, in order to project the virtual environment correctly. A dynamic stack can directly make use of that.

This simple interactivity plays into Gleicher's final point as listed above. If some visual fidelity can be sacrificed to bring many layers closer together, e. g. to compare larger features faster, a dynamic stack can accommodate this change faster and more effortlessly than the other layouts. Interactivity in the form of *linking* is also an advantage of layer stacks over grids: while a cursor or a highlight can be synchronized over all layers of a grid in the typical MCV fashion if the user moves it on one of them (this was implemented by Lobo et al.), a stack could also show a line directly connecting all layers at the place of the cursor, or at certain positions such as shown in Figure 4.2. Such a line is possible because from the user's point of view, the layers are only displaced vertically, and not in multiple directions as in a grid — it would be one straight, or slightly bent line, and not a network of rays that a grid would produce. Sequential scanning of these layers would thus also be in a straight line up and down, and not in a zigzag pattern, with much smaller changes to the visual field, and potentially less divided attention as a result.

With a stacked arrangement, Gleicher's concern about "good ordering" could also be better addressed: as can already be seen in some illustrations from Figure 4.1, layers could be sorted by e. g. complexity. Alternatively, layers that are closely related could be placed adjacent to further highlight their connections, such as the top two layers in the example from Figure 4.2. Their distance to each other would be minimal, thus resulting in a further decrease in eye movement necessary to compare them, and could alleviate Lobo et al.'s concerns regarding divided attention in juxtapositions.

A dynamic layer stack for as described here could be a first answer to the hypothesis from the previous chapter: immersive environments, through binocular presentation and their inherent awareness of the viewer's position, allow map layers to be tilted while staying legible, and thus are able to fit multiple layers in more efficient juxtaposed arrangements than possible otherwise, while offering useful interactivity. A more concrete hypothesis thus follows: *the dynamic layer stack could enhance multilayer map comparisons, compared to alternative methods such as regular juxtaposition of these layers in a grid, or temporal juxtaposition / blitting.* This new hypothesis will be tested in later parts of this thesis through an implementation and user study.

The dynamic layer stack is, however, not the only result from last chapter's more general hypothesis. As a solution to displaying multiple layers more efficiently, it by design uses much less space than is available in immersive environments, to keep divided attention to a minimum and maximize the perceived coherence between the different layers. The remaining space can however be utilized for other applications, not directly tackled by the layer stack as it is presented above, such as immersive visualization of urban data at *multiple scales*.

5 Data Layer Deformation: Immersive Focus + Context

5.1 Genesis: "Into the Layer Stack"

Initially, the layer stack (section 4.4) was imagined to be viewed not only from the above-described position that optimizes for an equal presentation of all layers, but also as an immersive environment itself. From the position of overviewing all layers, the user was envisioned to have the option to seamlessly transport into a first-person view on top of the lowest layer, with at least one other layer remaining visible above. All layers would be scaled up during this transition. This would result in enabling a more *focused* exploration of a layer that would benefit from a much closer perspective, like a layer containing 3D buildings, while the layer above could provide a *context*; either as an orientation help with e. g. street networks, or by showing larger-scale data, like urban zoning or meteorological information.

Street-level perspectives of 3D buildings are indispensable for e. g. qualitative visibility analysis in urban planning — Ortner et al. demonstrate an integrated system for such that relies on multiple coordinated views to show such perspectives from different positions and an overview from above [Ort+16]. Brasebin et al. argue that the "ability to be able to choose any viewpoint" in 3D views of urban scenes can be "crucial in the understanding of both physical organization of urban areas, their volumetric dimension and abstract data that might be presented along with the 3D built environment," especially in the context of public participation [Bra+16]. Expanding beyond the built environment, Graciano et al. propose a framework that visualizes volumetric data of not just cities but their terrain and its geological composition, arguing that "a proper 3D visualization method that shows those heterogeneous data together can be very useful for geoscientific and GIS professionals" [Gra+17].

Using the layer system as described above, these kinds of visualizations can be displayed as the "base layer," just in a perceptively larger version that is also stereoscopic. This leaves the sky above them, which is usually only decoratively used, if at all, as an additional canvas to project geospatial data for information or context. This approach could benefit immersive environments especially, where the visualizations can surround the viewer completely and glancing upwards is an intuitive interaction that does not require

additional input methods. A system like that could more fully utilize the large available space, as discussed in section 3.2. Similar to the previous chapter's initial hypothesis (section 4), this idea leads to the next one: *Immersive environments allow for novel methods of detail + overview, or focus + context visualizations of geospatial urban data.*

5.2 Related Work: Focus and Context

5.2.1 Benefits and Problems of 3D City Views

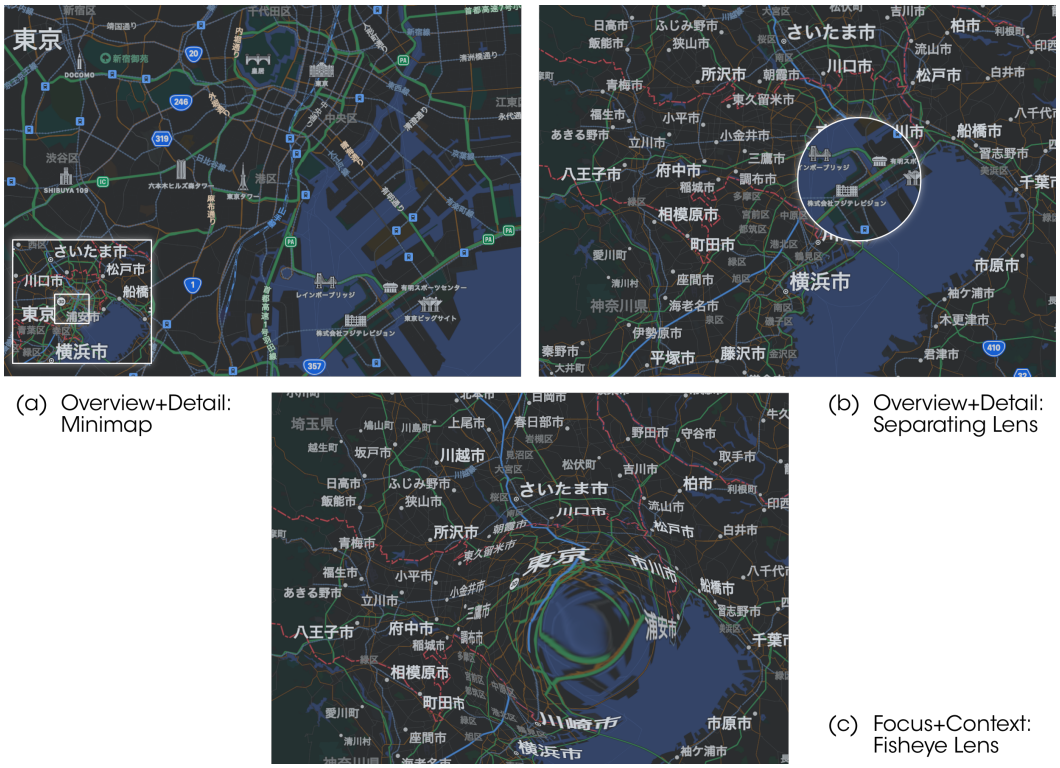
As already touched upon in section 2.3 and quickly reiterated above, there is large value in visualizing the full shape of cities and not just map projections for many applications. Representations that are as realistic as possible, or at least such that preserve the urban morphology as it can be seen by its inhabitants — 3D perspective drawings or renderings — have already been deemed by Kevin Lynch as essential in conveying a *sense of place* [Lyn60]. Such rather qualitative considerations are being carried over to the newest digital technologies, to which a chapter of Piga and Salerno's 2017 book *Urban Design and Representation* is dedicated, where the benefit of realistic perspectives for "ordinary citizens," stakeholders, and urbanists is explored for e. g. interpretation, assessment, and analysis of urban scenes and projects [Sal17]. Polys et al. further highlight the benefits of 3D modeling and immersive displays specifically in being beneficial as planning tools for decision-makers, compared to 2D presentations [Pol+18].

As beneficial for many purposes as they are, three-dimensional views of non-flat environments can introduce problems for visibility of the scenery. Geovirtual models such as buildings and hills will occlude anything behind them, such as texture, terrain, and other objects. Depending on the perspective, only a small number of buildings may be visible, and even then only partially if they overlap. Perspective also causes distortion in the form of foreshortening, shrinking distant objects until they become illegible visual clutter.

The choice of perspective from which a scene is viewed thus lets the user set the trade-off between the clear visibility of a flat 2D map presentation (when viewing from straight above at an infinite distance) and the realistic, but limited (through occlusion and distortion) 3D view at eye level. Moving from the former to the latter point of view leads to an effect that can be seen as separating the scenes into two areas: the *focus* area, which remains legible throughout; and the *context* area, rendered illegible by occlusion and distortion.

5.2.2 Overview+Detail and Focus+Context

This kind of interaction, albeit applied mostly to 2D views, is being described by Cockburn et al. in their review of interfaces for "working at and mov[ing] between focused and contextual views of a data set" [CKB09] as belonging to "zooming interfaces." Being dependent on user interaction to transition between the two views, it results in their *temporal* separation within a single display. As only one view is shown at a time, a certain burden is put on the user's short term memory "associated with assimilating distinct views of a system." Aside from freely navigable 3D environments, digital maps that allow zooming and panning also fall under the same category, as do basic desktop applications like most image browsers and document editors.



(a) Overview+Detail: Minimap

(b) Overview+Detail: Separating Lens

(c) Focus+Context: Fisheye Lens

Fig. 5.1: Mockups of the minimap (a) and lens (b) variants of overview+detail, and the fisheye variant of focus+context, as defined by [CKB09].

An "overview+detail" interface exchanges this *temporal* separation of focused and contextual views for a *spatial* one, by showing both at the same time, however each in its own "distinct presentation space." Typical examples of this technique can be found in video games as "mini maps," showing a "broad overview of the game's information space" in a smaller window, usually superimposed near a corner of the main (focused) view. Many online maps provide a similar functionality, usually showing an outline of the main view's scope on a much larger-scale map (Figure 5.1 a). As with zooming, the same principle is also being applied to a host of regular desktop applications, such as the thumbnails in slide show presentation programs, and overviews of long files in modern text and source code editors, where they are also referred to as "minimaps."

These overview+detail techniques are closely related to coordinated and multiple views [Rob07], which were already discussed above (section 4.3) in their applications for multiple layers of maps. Here, they are used for multiple scales of maps, though the representation of the overviews is often simplified to reduce visual complexity. A notable difference between those two applications is the relative sizes of their views: the views of multilayer systems are usually equally-sized, while most overview+detail systems show a large view of the focused view, with the overview being just thumbnail-sized, and off to the side or superimposed over a corner of the main view.

An inversion of this arrangement can be found when the lens metaphor is applied: a small, zoomed-in view superimposed over the place of interest of the overview (Figure 5.1 b). According to Cockburn et al. [CKB09], if the seam between the focused view and the context is removed, e. g. by using "fisheye" magnification, so that all parts are concurrently visible, focus and context become integrated in a technique called *focus+context*. As seen in the mockup of such a lens in Figure 5.1 (c), this coherent view of both focus and context comes at the price of introducing distortion, while potentially decreasing "short term memory load associated with assimilating distinct views of a system" and thereby improving "user ability to comprehend and manipulate the information," Cockburn et al. argue. With maps specifically, such distortions may however do more harm than good, as their shape is integral to their understanding. For this reason, Pindat et al. have proposed a technique that dynamically adapts the lens to the shape of the objects of interest, and have shown how their "JellyLenses" outperform regular fisheye lenses in multi-scale visual search tasks [Pin+12]. Another refinement of the lens metaphor is proposed by Zhao et al., where all deformation is *conformal*, i. e. it minimizes distortion of smaller shapes by exactly preserving local angles, thus trying to maintain legibility of features and the relationship between focus and context areas [Zha+12].

Lenses can of course also be applied to 3D environments, and not just 2D maps. Yang et al. demonstrated various methods of deforming 3D views through different shapes of projection planes, with example implementations of local magnifications of city scenes [YCB05]. Their approach however is not specific to 3D scenes, as the projection plane could be replaced with any other static scene (e. g. a map) for the same effect.

5.2.3 Techniques: Nonlinear Raytracing and Projection

More recent developments that are specifically utilizing 3D environments have been based on *nonlinear ray tracing* [Grö95], which itself is inspired by the effects of spacetime curvature from massive objects (gravitational lensing [Bar10]) and travel at relativistic speeds [HTW90]. By bending the camera rays around obstacles, the aforementioned problem of occlusion can be handled [Cui+10], and highly flexible focus+context visualizations can be created by effectively merging views from different perspectives into one



Fig. 5.2: Disocclusion of regions of interest through multiperspective cameras by Wu and Popescu [WP18].

final projection [WP16]. The same *multiperspective* camera approach has been applied to immersive environments for navigation and occlusion management [WP18], where predefined regions of interest are disoccluded (Figure 5.2), and to aid in exploration 3D scenes through automatic occlusion removal [Wan+19].

Similar to the multiple blending and floating lens approaches to visualizing multiple map layers discussed in section 4.2 and section 4.3, these curved ray cameras usually require careful configuration and setup. The other extreme on the scale of simplicity to providing concurrent focus+context has been suggested by Cockburn et al. in simply increasing "the size and resolution of the display, allowing natural human methods such as eye gaze direction, peripheral vision, and control of proximity to the work surface" [CKB09], which is yet another point in favor of large, high-resolution displays, and by extension immersive environments, as discussed previously in section 3.2.

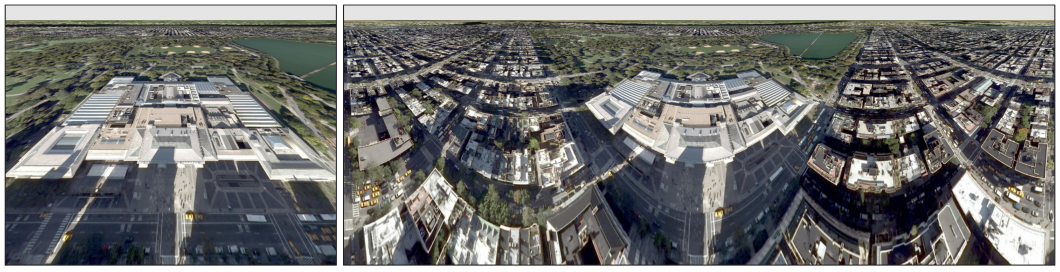


Fig. 5.3: Ordinary (left) and panoramic (right) perspective views of a city scene. The information beside and behind the virtual camera is lost in the ordinary perspective view, whereas it is preserved in the panoramic perspective view. In both perspective mappings, the information far away from the viewpoint is compressed into to a singular line, the horizon (from [Böt+08]).

More directed focus+context than just presenting a large and detailed map for the viewer to approach closely can also be achieved by global techniques, as opposed to selective or localized lenses and distortions. For flat maps without occluding 3D elements, the simple

perspective projection is already a kind of focus+context view, in which the focused area — that closest to the viewer — is displayed larger and in more detail, while the context gets progressively compressed to the point of illegibility and beyond with increased distance, up to a singularity at the horizon. Using a slightly more complicated projection with a cylinder as the surface, a panoramic view is possible that shows the context all around the viewer, and not just in one direction, as shown in Figure 5.3.

Böttger et al. explored transformations of aerial and satellite imagery that go beyond these perspective views in an attempt to overcome their limitations [Böt+08]. Beyond the infinite compression at the horizon, perspective projections compress shapes *anisotropically*, i. e. not equally for all directions, leading to "degenerated shapes" which become unrecognizable for large differences in scaling factors. Opposite of that are *conformal* projections, which minimize the distortion of smaller shapes by exactly preserving local angles, thus maintaining the legibility of features and the relationship between focus and context areas.

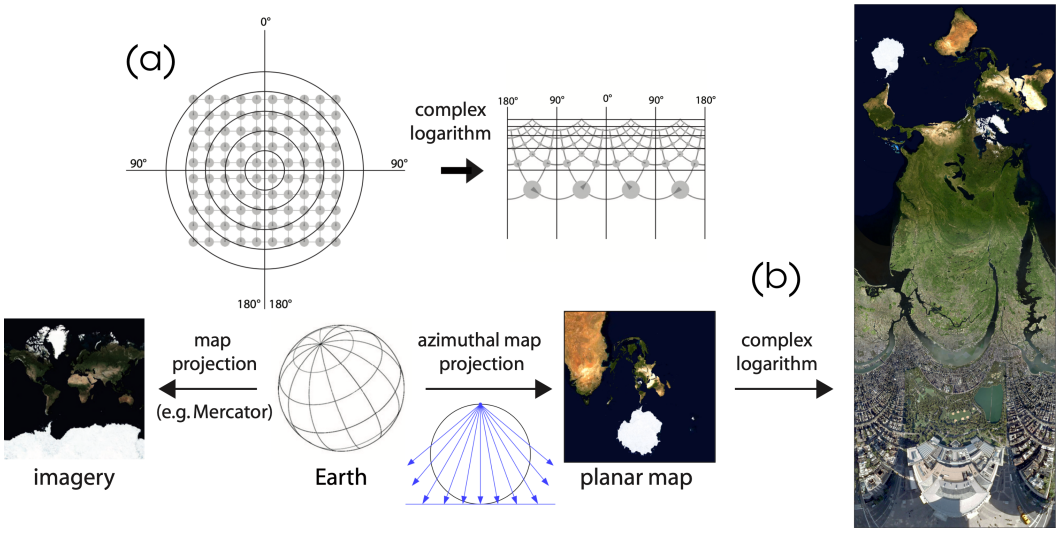


Fig. 5.4: The characteristics of the complex logarithmic transformation Böttger et al. employ to provide a focus+context view, and its application to show a city block in global context by using a planar projection of the whole Earth as its input (adapted from [Böt+08]).

The *complex logarithmic view* (Figure 5.4 a) that Böttger et al. employ to create a "detail-in-context" view appears at first similar to a panoramic perspective, but is in fact conformal: small circles are mapped to other small circles without local distortion or anisotropic compression. Concentric circles are however mapped to horizontals, directions to verticals, and a cut is introduced. Using a sufficiently large map as a source image, e. g. an azimuthal projection of the Earth's surface, the complex logarithmic view is able to visualize arbitrarily small focus views within the context of the whole globe, while maintaining a degree of legibility up to the farthest regions — as seen in Figure 5.4 b, Africa, Antarctica, Asia, Australia, Eurasia, and South America in the context view appear at the same size as a single building in the focus view, and all are still recognizable for what they are.

5.2.4 Techniques: Geometry Bending

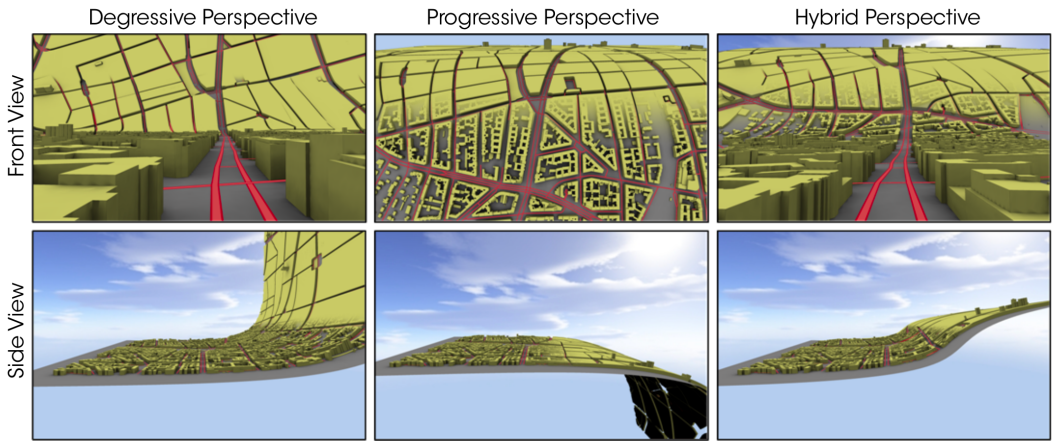


Fig. 5.5: Front and side views of the geometry-bending approach to focus+context, with different configurations (from [Pas+14])

A different approach to focus+context for maps is to deform the map or digital city model itself in 3D, instead of manipulating camera rays or using different projection schemes like the above examples. Figure 5.5 shows three variants of bending a 3D city model, shown from the intended (front) view, and the side; taken from the latest version of a system [Pas+14] that Lorenz, Pasewaldt, Trapp, and Döllner have been developing since 2008 [Lor+08; PTD11; PD13]. By defining a transition region in which the city's geometry is bent along a Bézier or B-spline curve and selecting the orientation of the focus region (before the bend) towards the camera, different perspectives can be merged into one view, with distortion occurring only inside the bend, aside from the perspective distortion inherent to a 3D view. A "pedestrian perspective" can thus be extended with a "bird's eye view" of the context region after the visibility drops beyond a few city blocks, and the horizon from a pedestrian perspective can be added to an initial bird's eye view (Figure 5.5 "degressive" and "progressive" perspectives). Bending the geometry twice, or making the spline more complex, "hybrid" perspectives can be created that combine pedestrian views for the focus area, bird's eye for the immediate context, and a horizon, which may be helpful for orientation if it contains landmarks such as tall buildings or mountains.

A very similar technique was presented by Möser et al. [Mös+08], which also includes localized conformal deformations to enlarge and highlight specific regions of interest, such as significant buildings, and disoccludes façades along a projected path using "adaptive iso-projection" (Figure 5.6). Both techniques can also incorporate mechanisms to decrease the visual complexity in the context views by blending into more abstracted cartographic styles or simplifying the city's geometry.

Another similar approach has been used by Veas et al. to create an "extended overview" [Vea+12], however implemented as two camera perspectives that are interpolated in the transition zone (Figure 5.7). In their application, the overview provided by the secondary

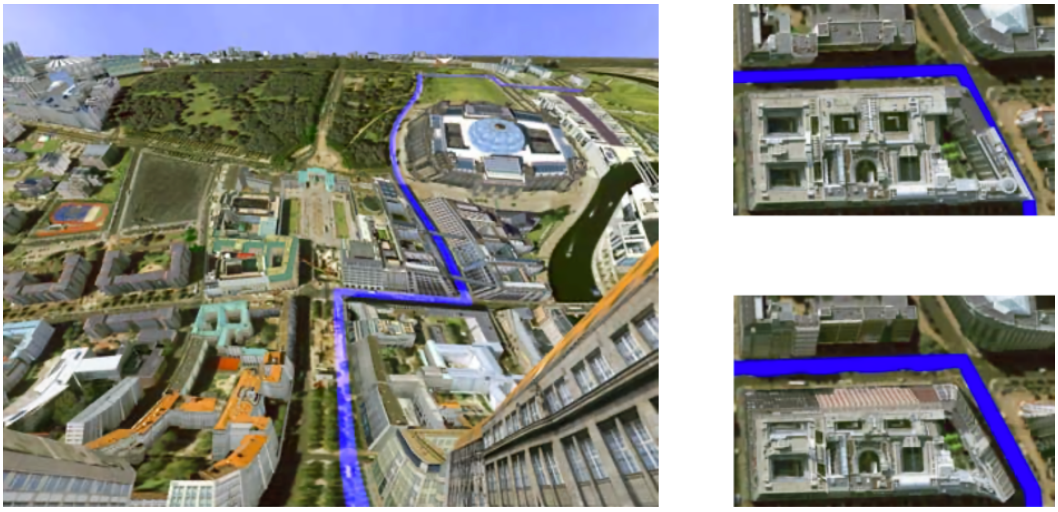


Fig. 5.6: Multiperspective focus+context view with local deformations of regions of interest, and disocclusion of façades (from [Mös+08]).

camera is rendered as transparent wireframes and overlaid on the image of a the primary *physical* camera view, thus augmenting the real-life focus with a virtual context, resulting in an AR view. A virtual primary camera view can also be overlaid over the physical one, adding information to buildings or showing contour lines on terrain.

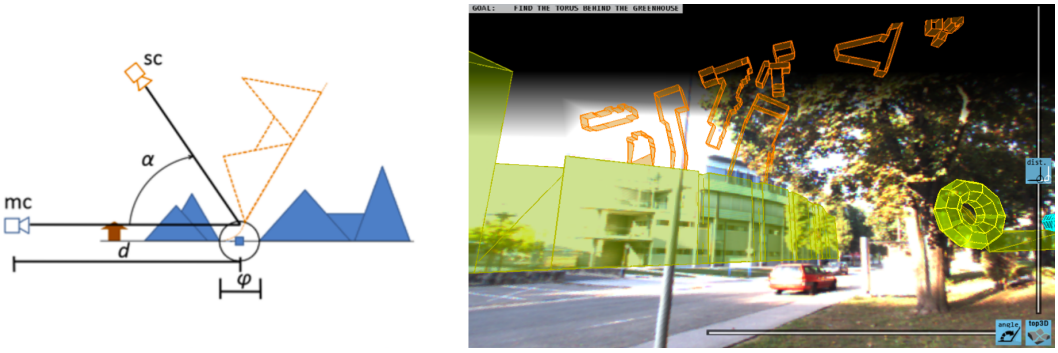


Fig. 5.7: The "extended overview" approach by Veas et al. to overlay a secondary camera's context view over an AR view (from [Vea+12]).

All three geometry-bending approaches above merge multiple perspectives into one view with minimal additional distortion outside their transition zones or bends, and are designed to work with 3D geometry, in contrast to the projection-based approach by Böttger et al., which introduces significant distortion in most of the resulting image. The trade-off however is a narrow perspective: the bending technique only reveals the context that lies behind the focus area, showing at most a narrow slice of the overall landscape, whereas the complex logarithmic view is fully panoramic and reveals the context in all directions.

This *directionality* is a major limitation: as the techniques all bend the terrain about one axis that is perpendicular to the viewing direction, their usefulness for context

information decreases with its distance from the viewing axis: with a wider field of view, the extreme left and right of the image would suffer a similar perspective distortion the bending tries to mitigate. In an interactive application, a rotation of the view by more than 90 degrees to either side would completely negate the bending's benefits if the bending axis remains static, or result in a rotated view with different distortions with every move if the axis was to follow the view direction, which could be disorienting. As Böttger et al. point out: "it is more difficult to recognize the shape of the object after a rotation, especially if there exists a clear notion of the object's orientation—for example, the shape of the European continent is often not recognized if it is presented upside down" [Böt+08].

This makes the bending approach ill-suited for interactive applications that intend to show more than one orientation, especially in immersive environments — Pasewaldt et al. place this limitation on their system: "the 3D panorama is not intended for virtual reality simulations but for analytic and exploratory information display," even observing simulator sickness in their tests with fast rotations [Pas+14]. They also point out a missing horizon (besides the "unusual depiction") as another source of disorientation and resulting simulator sickness, which would be less of an issue with the augmented overlay approach by Veas et al. in AR, where the real environment remains visible [Vea+12].

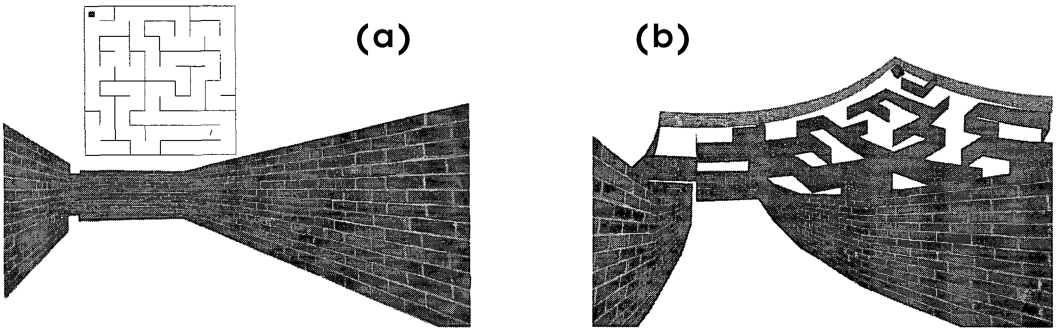


Fig. 5.8: A detail and overview of a maze (a), and Vallance and Calder's omnidirectional geometry bending approach to reveal context (b): "Data is distorted according to its distance from the user, so that close data is the same as a normal first person view, while distorted data approximates looking at that data from a second person viewpoint above the maze" (from [VC01]).

A compromise can be achieved by bending the geometry not along a straight axis, perpendicular to the viewing direction as in the above approaches, but along a circular one, centered on the viewer. This technique, using the same idea of having a bent transition zone after a certain distance from the viewer as in the works of Möser et al. [Mös+08] and Lorenz et al. [Lor+08] has already been demonstrated in 2001 by Vallance and Calder [VC01], but in the omnidirectional way described above, where the distortion is the same in every direction. As can be seen in Figure 5.8, while the revealed context is not planar as in the unidirectional bending approaches (straight lines become curved), it remains legible, and does not show the extreme distortions visible in the complex logarithmic approach by Böttger et al. [Böt+08].

Demonstrated for the application of navigating a maze, this technique was developed with interactive navigation and exploration in mind, unlike the unidirectional bending approaches which are intended for a single perspective. In fact, the authors specifically point out that GIS data would "lend itself well to navigation by bending," as it "tends to be very large and planar," and suggest augmented reality applications by overlaying a map "over the top of real scenery and then distorted in the distance to provide easy navigation." They argue that a technique like this could find application "anywhere the data is mainly on a single plane, and the data in the distance provides information necessary for navigation," especially when "close data needs to be viewed to facilitate local interaction."

Vallance and Calder list limitations of their technique that also apply to the unidirectional deformation schemes. The first is a "loss of context immediately around the viewpoint," which means that data just behind obscuring objects very close to the viewer may still be occluded, unless the focus area is made very small and the bend angle very steep. Conversely, the indiscriminate nature of the uniform environment bending approach may distort data that would be have been well visible from the first-person view without deformation. Information that is to be viewed without distortion thus must be close to the viewpoint.

Another limitation Vallance and Calder point out is the potential complexity of interacting with the parameters of the deformation: exposing these may "unnecessarily complicate the interface," even though changing them interactively could have drastic effects on the technique's efficacy. The authors suggest varying the bend threshold as an "effective way of controlling the amount of context," while other parameters such as the type and smoothness of the bend "may be less intuitive to manipulate." This of course applies even more so to the comparatively complex deformation used by Pasewaldt et al., which relies on parametrized b-splines with many control points that each contribute substantially to the result, as shown in Figure 5.9.

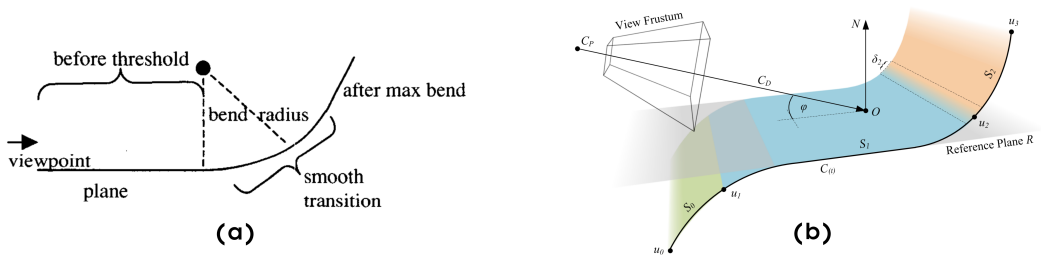


Fig. 5.9: The simple parametrization of Vallance and Calder's [VC01] geometry bending technique (a), and the more complex b-spline (b) used by Pasewaldt et al. [PTD11].

5.2.5 Synthesis and Design Considerations

This tour through related work on focus+context geovisualization with an urban focus presented a broad range of approaches and techniques. To choose which of these to use, develop further, or improve on, some design considerations need to be made for the intended attributes of a new system. The intention is to develop a technique that strikes a balance between preserving legibility of the urban form and reshaping its visualization to allow for focus+context. On a higher level, this means presenting professionals in urbanism as well as novice users with a legible and recognizable view of the city, which at the same time offers a new perspective on the relationship between a region in focus and its surrounding context, aiding in the visual analytics process, exploration, navigation in AR, or communication of urban development proposals. On a lower level, this encompasses the following specific requirements:

- **Present 3D city models in familiar perspectives:** In many urban applications, particularly in those that include participatory elements, a very common perspective is one that shows a region of the size of a city block from above the rooftops, giving visibility to the shape of individual buildings without too much occlusion [Bra+16], but lower perspectives, down to eye-level are also often seen [Pol+18]. Such 3D views have been compared favorably to 2D maps when applied to participatory planning projects, requiring less effort from users to understand and manipulate the views [Jac14]. These views should thus act as the *focus* views, without any distortion that could negate their aforementioned benefits.
- **Reveal context hidden by occlusions or perspective:** Objects that are outside the focus view, are obstructed by it, or rendered illegible by perspective should be made visible to place the focus view into its context. The resulting focus+context view should not introduce discontinuities, which would increase mental strain by making the user having to "mentally reconnect them" [Böt+08]. The visibility of connections (physical or conceptual) between close and distant regions should be enhanced, and the applied deformation should be minimally distorting.
- **Provide simple controls over focus+context:** The relative size of the focus area within the context and the amount of context made visible should be adjustable in real time, according to the needs of the user in exploring different views or finding an optimal balance. Control of the relevant parameters should be intuitive.
- **Benefit from immersive technologies:** Immersive environments can offer significant benefits over other display methods (section 3), and are increasingly being considered for use in urban analytics [Che+17]. By their nature, they provide an omnidirectional canvas on which to project data, as opposed to the limited windows traditional monitors present — a view of every direction is available through natural head movement by the user, which is ideally not restricted. The desired technique needs to work equally well in any direction the user looks, without having to change its characteristics to accommodate rotations.

These requirements eliminate approaches such as MCVs or similar overview+detail techniques, which would introduce discontinuities. The lens metaphor would only work in its inverse application, as the user should control how and how much of context is added, while the focus area remains unchanged from its original form.

For an implementation in immersive environments, a 2D projection like Böttger's complex logarithmic panoramas may not be ideal, and the complex setup and interactions required for nonlinear raytracing techniques may also be a hindrance to an intuitive application suitable for all stakeholders in urban projects. The problem of unidirectionality with the bending techniques by Möser et al. and Lorenz et al. for immersive exploration and navigation is addressed by the unidirectional bending proposed by Vallance and Calder, at the price of more severe distortion.

This inherent distortion in all focus+context approaches discussed above is another challenge specifically for applications in urbanism, where a representation or at least perspective of the urban morphology that is as close as possible to what can be seen in reality is highly desirable, and is instrumental in providing the viewer with a sense of place [Sal17]. Realistic representation of the city's shape is also integral message of Kevin Lynch's *Image of the City* [Lyn60]. A conformal distortion, as used by Böttger's projection scheme (Figure 5.4) that preserves local shapes could help to minimize this issue. Another possibility that would leave the first-person view completely intact is already suggested by Vallance and Calder [VC01], and shown in the approach of Veas et al. [Vea+12]: that of overlaying the distorted map "over the top of the real city" in an augmented reality application — the same could of course be done in virtual reality, with a realistic virtual city model staying intact and a distorted version of it being transparently superimposed for context and navigation. This approach also opens up the possibility of showing a different layer for the context view — either a simplified map, as achieved through blending by Lorenz et al. and Pasewaldt et al., or an additional layer of information, as discussed in the previous chapter on multilayer data (section 4).

Immersive environments thus offer the possibility to apply multiperspective focus+context geospatial visualization to real urban scenes via augmented reality, while virtual reality can be used to simulate these AR applications, and go beyond their constraints of position and interaction. The considerations made in previous sections about the advantages of immersive environments for urban data visualization apply even more to focus+context uses, where more of the large visual field can be utilized through and deformations.

The challenge now lies in designing a capable focus+context technique that can fully take advantage of the opportunities offered by immersive environments without being too complex to use and without distorting the city's image beyond recognition and usefulness. The following section describes a technique that aims to fulfill the above requirements: *inverse stereographic projection*.

5.3 Inverse Stereographic Projection for Multilayered Focus+Context

5.3.1 From Layer Stack to Focus+Context

Coming back to the layering system from section 4.4 as described at the beginning of this chapter (section 5.1) (zooming into its base layer with an additional layer floating above), compared to the related work in that state it can be classified as an overview+detail system at best. Functionally it most closely resembles a first-person view with an additional mini-imap, similar to the view in Figure 5.8 (a), but with the map being horizontal and above the ground instead of vertical and in front of the viewer, becoming a "mini" map only by perspective foreshortening through its distance. Looking up could aid in navigation or give additional data for the region, as it already can make occluded information visible, and improve visibility of that information by having a smaller viewing angle α than the original angle β, as shown in Figure 5.10 (a). This angle (the deviation from the ideal 90 degrees) gets smaller, the higher the layer floats. But as usual with detail+overview, this view is not connected to the focus view. In fact, if used as an additional data layer, viewing the information for the focus area requires the most effort (looking straight up). Establishing a mental connection between the floating layer and the ground could also be a challenge.

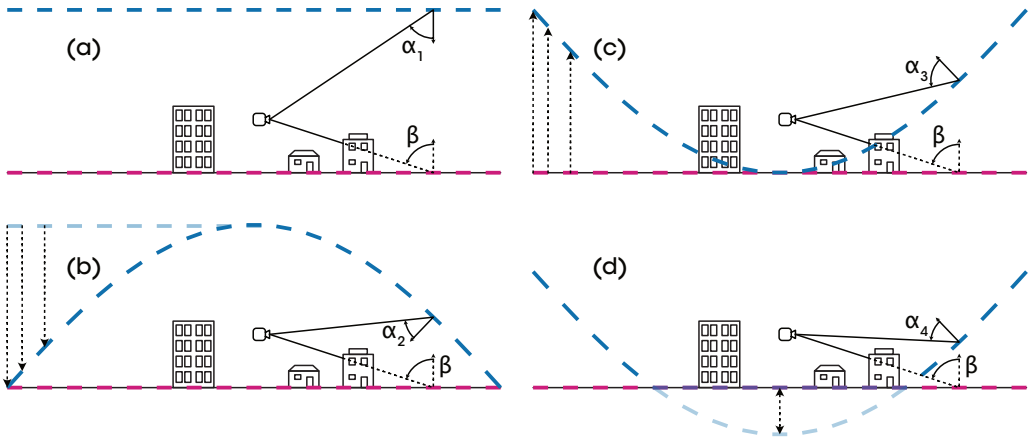


Fig. 5.10: Four iterations on the layered data approach from section 4.4: (a) the "sky layer," providing an overview over a first-person view, (b) the "dome," bending the layer down to establish a connection with the ground, (c) the "bowl," an inversion of the dome, creating a focus+context view, and (d) a flattened bowl, creating a larger focus area where the additional layer is directly overlaid on the ground. Viewing angles of distant objects in the additional layer are shown with $\alpha_{1..4}$ in comparison with the equivalent original object's β.

A first modification to the initial system (seen in Figure 5.10 a) was to bend the floating layer down, to establish a connection between the original view and the floating layer, forming a dome (Figure 5.10 b). Displacing the layer's contents only vertically, this con-

nection comes without introducing a discontinuity, but at the expense of stretching the geometry (seen in the varying length in the blue line segments in Figure 5.10), and hiding any context behind that connection. On the other hand, this deformation could enhance the visibility of far-away features further, as they are made larger by this stretching, and the viewing angle α can be even closer to the ideal 90 degrees, depending on the altitude of the viewer and on how steep the bend is.

Mirroring this dome around the ground plane results in a bowl-like shape, seen in Figure 5.10 c. With the bent layer now touching the ground just below the viewer and extending potentially infinitely into the sky, this scheme is now a focus+context technique, seamlessly revealing context around a unchanged focus area, with the possibility of also being a multi-layer technique if the bent layer is a transparent layer of additional information. The viewing angle α is still smaller than the original angle β on the ground, but to a lesser degree than in the dome scenario. The distance between original data and the additional layer is also reversed from the dome: it is minimal closest to the viewer, and increases into the distance. Common to both the dome and the "bowl" is the complete distortion of the context layer: any point that is not exactly above or below the viewer is being bent.

The next iteration on this scheme is creating an intersection between the "bowl" and the ground plane by translating the bent layer downwards, seen in Figure 5.10 d. Instead of making the ground layer transparent to see the additional layer through it, the "bowl" could be flattened between its intersection with the ground, and overlaid immediately on top of it, thus eliminating any distortion in the focus area (in the case of it being an additional data layer — if it is a duplicate of the ground data, it can be just omitted there). Lowering the bent layer also decreases the viewing angle α again, however occlusion through objects in the focus area can become a bigger issue than in the previous schemes (Figure 5.10 a–c).

This last scheme most closely resembles the omnidirectional bending approach by Vallance and Calder [VC01], with the following differences:

- The **transition** between focus and context is not a smooth curve, but an edge, the sharpness of which depends the bending parameters. This could help with the issue of misinterpreting the bend for "as a mountainous region," as Pasewaldt et al. have encountered with their technique ([Pas+14]).
- The **deformation** is a stretching of the geometry by pulling it down from above/up from the ground, instead of applying a bend within a defined transition zone. Distant objects are thus not pulled closer to the viewer, and are distorted in a different way.

- The bent layer can provide **additional data** on the focus area by being transparently overlaid on it, in addition as acting as context or a navigational assistance in the distance as suggested by Vallance and Calder.

These schemes, as well as Vallance's and Calder's technique distort the geometry to different degrees depending on the size of the transition zone, and the angle of the bend, or the choice of curve the layer is being bent to. The resulting shape is quite arbitrary, and parameterizing it for real-time use could be challenging, as the authors have pointed out [VC01]. This initial, simple attempt at omnidirectional focus+context with the flattened "bowl" (Figure 5.10 d) already shows some potential improvements over previous techniques (clearer transition, additional layer), but also shares mostly the same downsides, another one of which not yet mentioned is their non-conformal distortion of geometry. Böttger et al. have already shown a technique that does a focus+context distortion in a conformal way, using a technique akin to cartographic projections [Böt+08]. Could something similar work for a three-dimensional, immersive environment?

5.3.2 The Sphere — an Ideal Shape?

It could be argued that the least "arbitrary" shape is the sphere, having only one parameter, its radius, and being the one most commonly occurring in nature at all scales, from vapor droplets, over to planets and stars, to globular clusters and cosmic supervoids. The most correct of the common representations of the Earth's surface features is a spherical globe, requiring the least amount of deformation from our planet's oblate spheroid shape. Naturally, the familiar shape of the physical globe translates well to virtual environments, being the preferred choice for world-wide geographic maps in a study conducted by Yang et al. [Yan+18b].

In addition to the many historical, societal, and psychological connotations of globes and spheres that are summed up by Ingold in an anthropological view of environmentalism [Ing93], they also possess qualities that make them well-suited for information visualization that goes beyond direct representations of other spherical objects. Brath and MacMurchy list the fact that a sphere wraps around itself with no boundaries as beneficial to "gestalt association of relationships between nearby objects," and the possibilities of intuitive navigation as advantages to using spheres for information visualization, alongside pleasing aesthetics of circular and spherical layouts [BM12]. The authors discuss examples of sphere-based visualizations of treemaps [HHL04] and graphs [Spr+97; Mun98], the latter of which formed a part of Tamara Munzner's dissertation on interactive visualization [Mun00].

Early methods for interacting with digital 3D environments employing spheres were the *Virtual Sphere* [CMS88] and *Arcball* [Sho92], which still find applications in many

modern applications such as Unity 3D, as reported in a recent survey on 3D virtual object manipulation [Men+19]. In immersive environments, spheres have also been used to create novel locomotion methods, such as the quite self-explaining *Floating Hamster Ball* by Hurtado et al., which adopts the "intuitive physics of rolling a sphere on the ground" [Hur+18a].

A physical precursor for spherical VR navigation, the *Virtusphere* [MFW08], being a "spherical cage that both traps and liberates users by projecting infinite environments at a one-to-one scale," is seen as an allegory of simulation itself by Canizares, while referencing to the ideal of DaVinci's Vitruvian Man, the purity of Buckminster Fuller's tensegrity spheres, and the equation of spheres to perfection throughout history [Can19]. Contemporary cultural critics Hito Steyerl and Vivian Sobchack [Sob16] are also quoted when drawing parallels between the Virtusphere and social media and economic bubbles, the modern ubiquity of screens, and other "spheres of meaning." Canizares finishes:

> "The powerful image of the Virtusphere is therefore less a representation of an object, and more an idealized diagram of an evolving subjectivity and relationship to alternate realities in general, with the self in the center surrounded by a layer of world—whatever that may be."

In this spirit, the sphere does seem like the ideal shape to project additional layers of information about the surrounding world, or to use it to reshape that world to better understand one's place within its context. A significant difference between the problem at hand — visualizing urban focus+context — and the above deliberations on the sphere being a canvas to project real and fictional worlds as well as abstract information onto, is the issue of projection. Both a virtual viewsphere and the almost-spherical Earth can easily be projected directly onto an egocentric sphere like the Virtusphere, while the layouts of abstract graphs and treemaps are mostly arbitrary anyway. There is alas no one "correct" way to map a flat surface, such as a city map onto a sphere.

In a study that compares 2D representations of graph visualizations to their equivalents in immersive environments, Kwon et al. needed techniques that did exactly that, however, and arrived at using the inverse of techniques that have been utilized in cartography for projecting the Earth's surface to flat maps [Kwo+16]. As there exist "literally hundreds of map projections for transforming spherical or ellipsoidal coordinates to Euclidean space" [Jen12], the choice has to be narrowed down. Of the conformal projections (those which preserve local shapes and angles) the *stereographic projection* is the one that seems most suitable for a focus+context application, as it is the only one that uses azimuthal perspective projection (projecting a spherical surface by extending lines from a point of perspective to a plane) to map a sphere to an infinite plane.

5.3.3 The Stereographic Projection

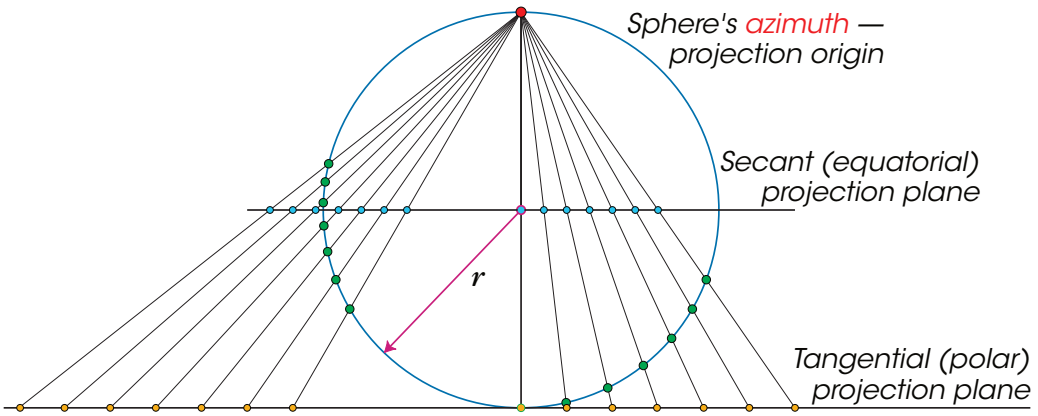


Fig. 5.11: Stereographic projection of points on a spherical surface to secant or tangent planes. Images on parallel planes differ only by an isometric scaling factor.

In its application for "unfolding of a spherical surface," the stereographic projection (Figure 5.11) was first described by Hipparchus during the second half of the second century BCE, who used it to map stars from the "celestial sphere" [Neu49]. Being a *geometric azimuthal* projection [DA34] means it can easily be constructed mechanically by drawing lines from one pole of the sphere towards a plane tangent to the opposite pole or one intersecting the equator — these lines project points on the sphere's surface to points on the plane (see Figure 5.11). As already stated, it is conformal: the shapes of "small" features appear correctly, or, more precisely, that relative angles at each point are kept [Sny87]. Great circles on the sphere are projected to spheres on a plane and vice versa. Larger and more irregular structures will appear deformed, though, as the projections of features on the plane are more compressed with increasing distance from the sphere. The choice of placing the plane at the opposite pole or the equator only results in an isometric scaling: Figure 5.11 shows how projected points on the secant plane are just as equally spaced as those on the tangential plane, only more compressed.

Those characteristics made possible the construction of the *planispheric astrolabe*, a tool that used a star map for navigation, and whose history goes back to at least Ptolemy [Neu49]. It took more than a millennium however for the stereographic projection to be used to map the earth, and not the stars, with the earliest-known example appearing in a 1507 world map by Gautier Lud [Sny87]. Since then, the stereographic projection has also found great use in geology and specifically crystallography, as first developed by William Hallowes Miller in 1839 [How96], and also in photography to create panoramic projections [Ger+07]. As a cartographic tool, it remains particularly useful in meteorological applications, where it is used to track moving weather features such as cyclones. The pole (antipodal to the projection origin) is centered and moved with the tracked feature, ensuring consistent distortion around the center regardless of absolute location [She+71]. These examples show that the characteristics of this projection were not only

found useful in many applications, but that this widespread use keeps familiarizing users with it, particularly within geospatial domains of expertise.

One downside for cartographic use of this projection is its inability to map a full sphere to a finite plane: as is seen in Figure 5.11, points closer to the projection origin on the sphere's pole will be mapped further and further out on the plane, approaching infinity as they approach that pole, with projected features being increasingly stretched out. For this reason, projections of the Earth rarely extend much further than one hemisphere. However as in the case of Kwon et al.'s inverse projection of graphs from a flat surfaces to a sphere centered on the viewer [Kwo+16], this limitation is actually an advantage for a focus+context application. The graphs visualized by Kwon et al. were finite in extent on a flat plane, as is the extent of the context region relevant to urban visualization. This being the case, both can be fully mapped on just a portion of a sphere, which is beneficial in two ways: the extreme distortions that occur near the projection origin are avoided, and the user can view the projected graph or context without exceeding a comfortable range of motion [You+92], depending on the relation between the sphere's radius and the extent of the plane.

5.3.4 Inverse Stereographic Projection for Focus+Context

If the tangential plane in Figure 5.11 represents the ground, and a viewer is located inside the sphere along the top-bottom axis, the inverse of the stereographic projection (plane to sphere) can already produce a result similar to the "bowl" shown in Figure 5.10 (c), with the notable differences being the spherical shape of the projected context region and a compression of it — where the bowl would extend indefinitely into the sky, the sphere wraps around, and could map an infinite extent of the ground plane on its finite surface. Conceptually, the projection origin being directly above the viewer, the sphere's surface that is between the viewer and the original environment acts as a screen that shows a compressed view from an elevated perspective, as long as the viewer is at any point below that origin. Starting from the point directly below the viewer, the image on that spherical screen gradually diverges from what the viewer would see directly of the ground below, showing instead what lies further ahead. This could already be seen as a gradual multiperspective focus+context view, insofar as the ground closest to the viewer is minimally distorted (the focus), and much more of the environment is visible than otherwise possible without the stereographic projection.

If the ground plane intersects the sphere between its pole and equator, the same projection mechanism will produce a result that shares the benefits discussed for the "flattened bowl" from Figure 5.10 (d): a true focus area with zero distortions, and a clear transition into a compressed context. The diameter of the sphere, and how it intersects the ground

(from being tangential to being bisected at the equator) are now the only two parameters which fully dictate the extent of the focus area, the viewing angle of the context, how much that context is being compressed, and how close to the viewer this compressed context is being projected.

This inverse stereographic projection from a secant plane by itself is already a novel idea, sharing all potential benefits of the "flattened bowl" over other techniques as discussed above, and overcoming its shortcomings of having no well-defined shape and parameterization, and being a conformal [1] projection. An additional benefit over any of the previously mentioned technique could be that the sphere actually pulls the projected context closer to the viewer as soon as it passes over its equator, an effect which could be perceived even more strongly in immersive environments with stereoscopic presentation. Finally, the sphere itself could act as a target for navigating and interacting with the environment, including being able to be used in ways similar to the virtual hamster ball by Hurtado et al. [Hur+18a].

The deliberations regarding the deformations show in Figure 5.10 did omit a discussion about how to handle height information, i. e. any geometry that protrudes from the ground plane, e. g. buildings and uneven terrain itself. The techniques that bend the environment around an axis, such as those by Vallance and Calder [VC01] or Lorentz et al. [Lor+08] have no problem with that: the whole 3D model gets bent, as shown in an illustration of the approach by Veas et al. in Figure 5.7. No part of the geometry is deformed when using a straight axis outside of the transition zones, where tall structures might get compressed or expanded, as seen in the results from the implementation by Möser et al. in Figure 5.6; the environment appears normal, just at a different angle. The method by which the bowl or dome deform the environment — pulling geometry up or down, depending on distance to the viewer — could be equally applied to 3D structures, which would e. g. lead walls appearing to rise from the ground at wrong angles. A compensating scheme could be developed to rectify the angles, but the underlying shapes are already distorted.

For the inverse stereographic projection to handle terrain elevation and buildings, its principle of connecting a point on a plane with the projection origin on the pole of a sphere could be extended to *any* point outside the sphere, as shown in Figure 5.12 a. This would require an inclusion of the height of a point into the projection's calculation, instead of choosing from one of the two existing formulas that require all points to lie on either the tangential or equatorial plane. Unlike the bending methods, this projection method will map all geometry to the surface of the sphere, with nothing protruding from it, stereoscopic depth perception is thus no longer possible for the projected context.

[1] Strictly speaking, the viewer would only really see a conformal projection of shapes from the plane if viewing the sphere exactly from its center, but as discussed in regard to the tilted layer stack approach in section 4.4, the perception of shapes is relatively tolerant towards being tilted [VGB05].

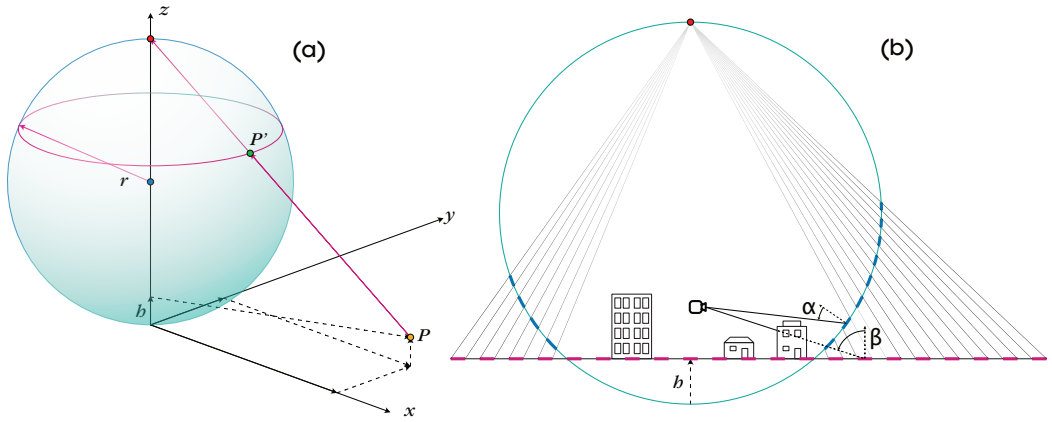


Fig. 5.12: Adaptation of the inverse stereographic principle to project any point P in 3D space to a point P' on the surface of a sphere (a), and in an application for urban focus+context (b). The sphere intersects the digital city model as it is elevated by h. Projected and original viewing angles α and β are compared as in Figure 5.10.

This results in the viewer inside the sphere seeing the full 3D perspective from the projection origin above, compressed on the sphere's surface. Any geometry that lies inside the sphere would be the unmodified focus area, similar to the "flattened bowl" from Figure 5.10 d, and the transition to projected geometry would be a clear edge at the intersection between terrain and sphere, shown in Figure 5.12 b. Compared to the flattened bowl, the viewing angle α is now even more reduced for a sphere that intersects the city in the same way. This technique now shows the perspective "from above" projected onto the sphere's surface, which appears compressed when viewed from any point below the projection origin, and thus also shows differences in elevation as seen from that view above.

5.3.5 Terrain Distortions and Flattening

Projections of slopes and vertical discontinuities on any non-flat terrain are however not conformal anymore: the conformity inherent to the classical stereographic projection gets disturbed wherever there are differences in terrain elevation. Changes in elevation slopes result in changes to the angles of the lines connecting points on those slopes — they are no longer coplanar with the connecting lines of other points, meaning they do not share the same distortion characteristics as for the planar case, and therefore do not preserve angles as lines on coplanar shapes do when projected. The projection still is "conformal" in the strictest sense only for infinitesimal distances, but not anymore when comparing the projected image on the sphere to a flat map of the same region.

The projection of the perspective "from above" also results in an exaggeration of elevation differences with an increased distance, where the angle from the projection origin gets closer to being horizontal. This increasing the influence of the vertical component of the original point compared to its horizontal over the projected point's vertical component.

Slopes facing away from the projection origin get compressed, and those facing toward it get expanded, as shown in Figure 5.13. As this origin is not infinitely high, sufficiently tall buildings or mountains may still cause occlusions even from that elevated perspective, which would be projected as an inversion if the projection consists of e. g. contour lines or sparse road networks, or in an overlap if projecting transparent textures (also shown in Figure 5.13).

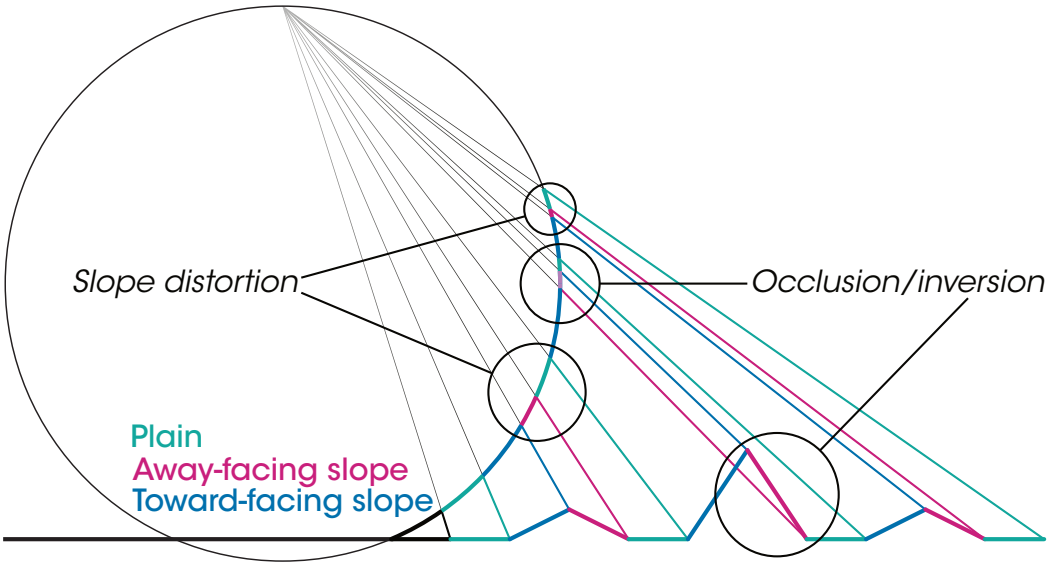


Fig. 5.13: Additional distortions in the inverse stereographic projection caused by uneven terrain: toward-facing slopes are expanded, while those facing away from the projection origin get compressed. Sufficiently high peaks can cause an inversion, where the away-facing slope and plains behind it may be occluded, or appear inverted if projected transparently.

These effects can be mitigated by a sphere that is larger in diameter to raise the projection origin high enough to prevent inversions and too much slope compression, but that may compromise its usability for focus+context in some cases by not raising close-by context high enough to be effective. It can at thus appear as though the effects from terrain distortion may be overly detrimental to the performance of the inverse stereographic projection for anything but flat city maps, where buildings only appear inside the focus view, or only very gently sloping terrain that does not require a large extent of context to be projected. However, this exaggeration of elevation could also be seen as beneficial in two different ways:

- First, it is not uncommon in cartography and geovisualization to enhance the height of terrain for emphasis, either by exaggerated relief shading in topographic maps, or by vertically scaling up terrain features for 3D visualization, especially when showing large scale regions or the whole planet, as a to-scale representation would make even tall mountains too small to easily see; the height of Mt. Everest being only 0.14% of the radius of the Earth. With this exaggeration increasing with

distance to the viewer, the stereographic projection can provide the same effect automatically and gradually.

- Second, traversing a hilly or mountainous region with the effects from seeing a projected and compressed view from above would make the context change in appearance more drastically than a flat environment. This movement would add to the already exaggerated terrain, resulting in a "kinetic depth effect," or *structure from motion* [VKB14], which can restore a sense of depth that is otherwise diminished through the "flat" projection onto the sphere's surface. Buddeberg et al. have found that vertical exaggeration and interactive animations that create such effects can significantly enhance the perception of differences in elevation in a user study with cartographers and visualization experts [BJW17].

Still, there may be cases where the above effects are undesirable — when the distortions are too strong, or when the projected context needs to be as true to a regular 2D map of it as possible, e. g. for navigation with a street map — but the focus area should still have the full 3D terrain. In those cases, the original context outside the sphere must be flat on one plane so that their projection on the sphere's surface remains conformal, without any of the distortions shown in Figure 5.13. This can be accomplished by pulling all geometry that lies outside the sphere vertically towards a secant plane, ideally one coplanar with the ground below the viewer. This flattened terrain is only used as a source for projection onto the sphere, and is otherwise not seen — in the case of the projected context acting as a transparent overlay, the original terrain is thus not visibly affected.

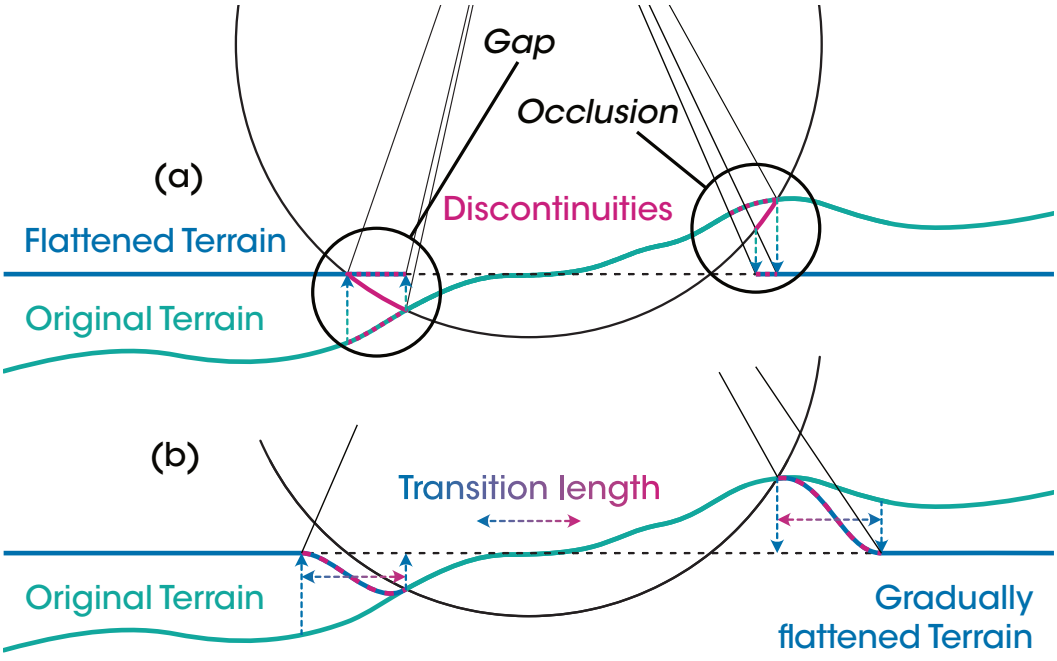


Fig. 5.14: The effects of flattening terrain outside the sphere for conformal stereographic projection (a), and a prevention of these possible gaps and occlusions by gradually flattening the terrain in a transition zone outside the sphere (b).

Figure 5.14 (a) shows how such flattening will result in gaps in the projection where the original terrain is below the reference plane, and occlusions where it is above, when any geometry outside the sphere is immediately flattened to that plane — discontinuities are introduced. These can be avoided by instating a transition zone, within which outside terrain will gradually morph towards the secant plane — not unlike other methods that generate smooth focus+context [PA08]. The extent of that zone can be chosen to be wide enough so that no inversions or occlusions occur as in Figure 5.13. The result is a well-defined, three-dimensional focus region free from distortion, gradually transitioning into a conformal projection which shows a compressed view of the context, as it would be seen from an infinitely high point, or on a map.

Of course, a compromise between this conformal projection and the potential benefits from the terrain exaggeration discussed above can also be achieved by only partly morphing the original terrain towards the plane — a parameter (from fully flat to no flattening) that can be interactively used by the viewer, or pre-set depending on use case. Even with adjustable flattening, the whole projection can still be controlled with only three parameters: the sphere's radius, the height at which it intersects the reference plane, and the degree of flattening. With that, there may be room for another parameter: the shape of the projection surface itself.

5.3.6 Other Projection Surfaces

Using the same principle of connecting a projection origin with points on the original surface, and imaging them to where these beams intersect with the projection surface, geometric shapes other than the sphere can become candidates to provide these surfaces. In cartographic projections, cones and cylinders have also found widespread use as targets for a variety of projections, where their characteristics make them more or less suitable depending on the desired application, e. g. which parts of the globe are to be shown, and which kinds of distortion are permissible [Sny87]. Being radially symmetric, basic shapes, both could be easily adapted for projection of context, as shown in Figure 5.15.

If only vertically upright positions for these shapes are considered (the only way to guarantee the same radial symmetry, and thus omnidirectionality of the projected context as the sphere provides), parameterization falls to the radius and height of the projection origin in the case of the cylinder, and also an angle at which the reference plane[2] is intersected in the case of the cone.

[2] The horizontal plane intersecting the ground below the viewer in case of uneven terrain.

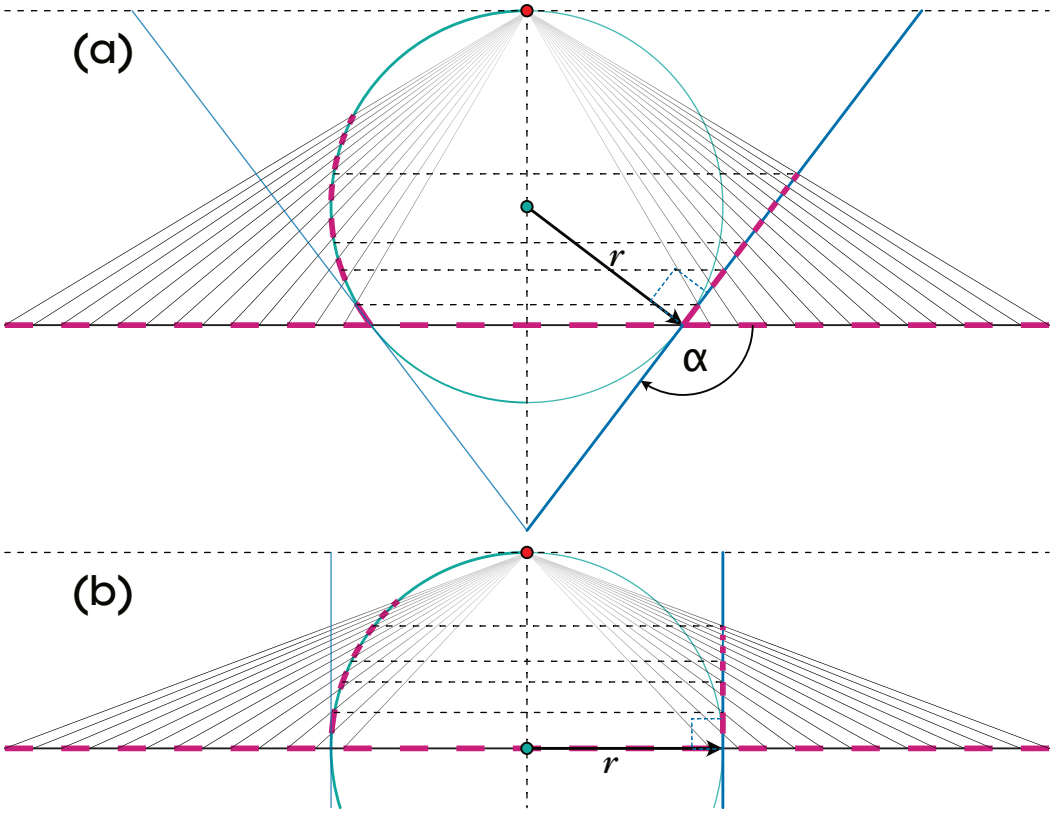


Fig. 5.15: The construction of conical (a) and cylindrical (b) projection surfaces from an "equivalent" sphere, by placing tangential surfaces at the intersection between sphere and the ground plane. Any point projected on the cone or cylinder (shown on the right half) is lower than on its equivalent sphere (shown on the left half), as indicated by dashed horizontal lines.

Another constraint can be made to further simplify the construction of these shapes and their parameterization by basing them on the sphere, the principle of which is shown in Figure 5.15 for both the cone and the cylinder. Taking any configuration of the inverse stereographic projection (section 5.3.4) — a sphere with a given radius that intersects a reference plane — a cone is constructed which intersects that plane in the same circle, and is tangential to the sphere at that circle. With that, the sphere's projection origin — its azimuth — can also be adopted, thus making the number of parameters needed to construct a cone for context projection the same as that of the sphere. A cylinder can then be created by having the sphere intersect the reference plane at its equator (compare Figure 5.15 a and b).

This constraint, or rather simplification, also has the effect of starting the context projection at the intersection in the same way as the sphere, only gradually diverging toward the sky: As can be seen in Figure 5.15, the cone and the cylinder (which can be seen as a special case of the cone, which intersects the plane at a right angle) compress the context faster; the same data is projected lower on their surface than would be on the "equivalent" sphere. Their projections are no longer conformal because of the additional vertical

compression compared to the sphere, and rather share characteristics with perspective projections shown in Figure 5.3 in that they have a horizon, at which the context becomes infinitely compressed (as opposed to the single point on the sphere) which leaves a gap in the sky where nothing is projected. The projected context extends away from the viewer until that rim (which is at the same height as the equivalent sphere's azimuth would be), appearing smaller, instead of wrapping back and enveloping as would be the case with the sphere, appearing magnified.

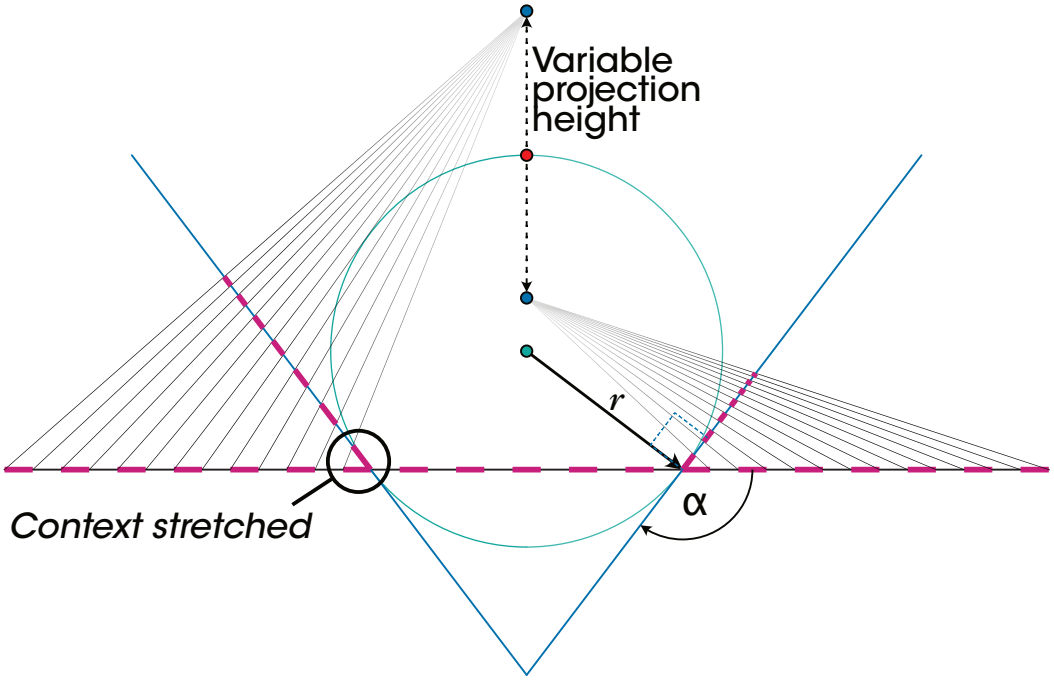


Fig. 5.16: Two variations of the projection origin's height shown for the cone (blue), compared to the original on the equivalent sphere's azimuth (red). If this height is increased, the projected context just behind the transition is stretched longer than it is originally.

This compression behavior can be altered further by decoupling the height of the projection origin from where it would be on their equivalent spheres, as shown in Figure 5.16. Setting it higher or lower than the original point stretches or compresses the projection, respectively, also making the transition from focus to context less smooth. In the case of stretching, the context may actually appear larger in the projection than it originally is just after the transition, before being compressed. This control over the compression behavior can make the cone and cylinder useful for scenarios where the transition rim lies directly behind occluding objects: the initial stretching from using a higher projection height would be hidden, and more of the context would be visible, as the horizon also rises with the projection height.

Besides this potential benefit, the projection characteristics of cones and spheres could be beneficial particularly to urban scenes with grids and straight lines, such as road networks. Whereas the inverse stereographic projection turns (infinite) straight lines from the context region into circles on the sphere, they become conic sections on the

other two shapes. Depending on the intersection angle and the position of the lines, they can form hyperbolas which may almost appear straight. The trade-off for this added flexibility is one additional parameter for the user, for a total of three: the radius and vertical position of the equivalent sphere, plus the new parameter for adjusted compression.

5.3.7 Double Stereographic Projection

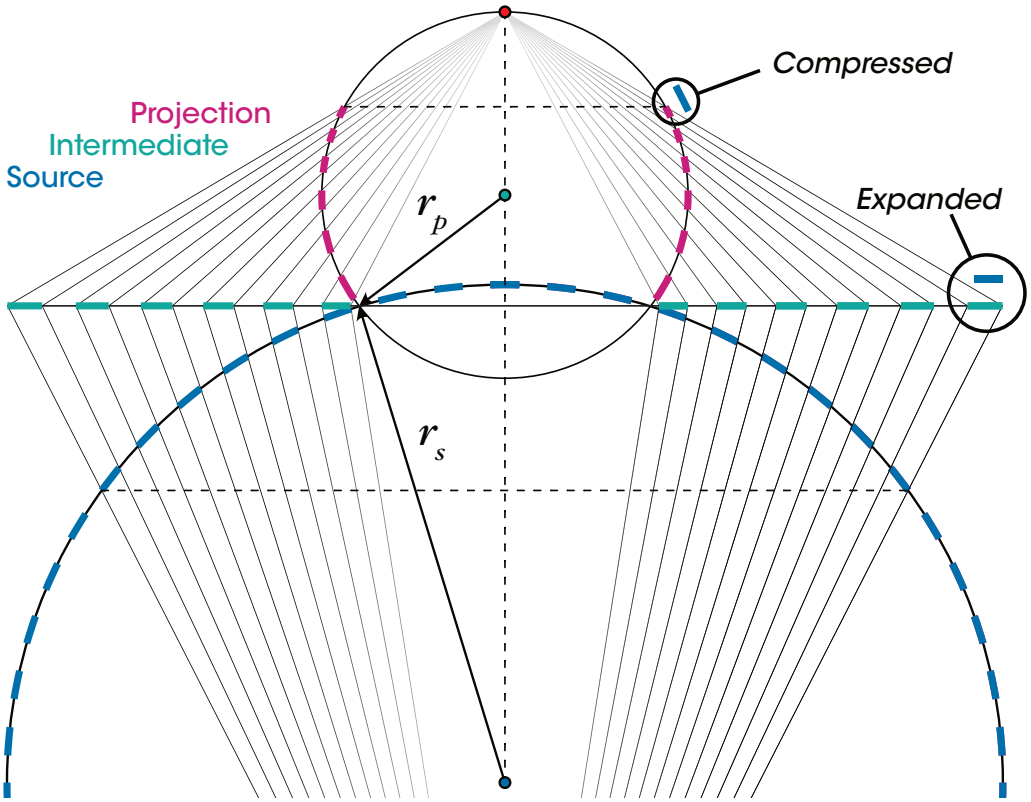


Fig. 5.17: Using the stereographic projection first on the surface of a whole planetary body (the source) to an intermediary plane, then a second time on the virtual sphere focus+context projection. The uniform features on the source body (blue) are used to illustrate the effects of each projection regarding expansion and compression: while the final projection is conformal to the source, it is non-uniformally compressed if its radius is smaller than the radius of the source body, or expanded if larger.

For geovisualizations on a scale where the curvature of the planetary body is non-negligible and relevant, the context is no longer the kind of "terrain" discussed above (section 5.3.5), i. e. consisting of fluctuations in elevation that can be flattened out if they cause unwanted distortion. The context is then also mostly spherical, in which case the stereographic projection could be applied twice: first to project the planet to an intersecting plane, and then this plane to the projection sphere, as shown in Figure 5.17.

The effect is distinct from simply viewing the planet as an inside-out globe as investigated by Yang et al. [Yan+18b], as there still exist an unmodified focus region (the dome from

the original planet above the intersecting plane), from which the context emerges and is lifted upwards by the second projection, which will show a compressed or expanded view, depending on the relationship between the size of the projection sphere and the original sphere. If both projections are the conformal stereographic projection as described in section 5.3.4, their concatenation will also remain conformal[3]. In fact, as the projections work against each other in a way (the first one expanding the data, and the second one compressing it), the result may appear much less distorted than the single stereographic projection from an intersecting plane. If the radius of the projecting sphere r_p is chosen to be the same as that of the source sphere r_s, the result will appear as a distortion-free inversion of the original planet.

For that reason, this approach may also be advantageous for geovisualization where the view of the curvature is not necessarily needed. The generation of this intermediary projection, which in effect is a kind of optimized map projection specifically for this use [4], is possible whenever a large part of the planetary body is mapped, and does not depend on specific ratios between the source and projection spheres. The result will in any case be a more gradual compression, and a completely projected sphere without a rim or a point in which the compression is infinite: the infinite expansion from the first stereographic projection is perfectly canceled out by the compression of the second one.

5.4 Conclusion

Geometric deformation is a powerful tool to go beyond detail+overview into focus+context visualizations, both for scientific and information visualization, and for urban and geovisualization. Immersive environments allow three-dimensional geometry in general, and viewer-centered focus+context specifically to be better perceived and interacted with than would be possible in desktop environments.

From regarding related work, the inverse stereographic projection with the enhancements discussed in this chapter (terrain flattening, different projection surfaces, planetary "double" projection) appears like a logical next step in investigating how the advantages immersive environments provide can be exploited for urban data visualization and beyond, as the sphere that is key to this projection is central to much of the field of immersion and perception, and the technique itself is a straightforward continuation of cartographic principles from antiquity.

[3] As defined by a projection that transforms circles into other circles, concatenations of conformal projections must also be conformal.

[4] This is similar to how Böttger et al. [Böt+08] fed their complex logarithmic projection with a variety of different projections from the globe, shown with the stereographic projection in Figure 5.4.

In both this chapter and the previous one, physical separation is the essence of the proposed techniques. Separating different layers of geospatial data into an actual stack of multiple and coordinated views aims to preserve a natural coherence while splitting information thematically, and the inverse stereographic projection gradually separates information from its original position where it may be obscured or otherwise illegible, using a method that strives to be as natural as possible. By gradually projecting the environment to the sky as it gets more distant and harder to see, a similar kind of coherence to its origin is intended as is with the MCV stack. The transition can be smooth and clear at the same time, or can be emphatically abrupt to prevent misinterpretation as to its distortion of the environment, depending on its configuration, and can act as a kind of inverse lens through which the continuation of the focus area can be seen far into the distance, as well as an overlay of additional information by using transparent or sparse layers for the projection while keeping the environment otherwise intact.

Both the stack and the inverse stereographic projection are intended for use in immersive environments first: better depth perception and a large virtual workspace, up to one that can fully surround the viewer are essential to their utility. While the difference between augmented and virtual reality is not essential to the MCV stack, as it merely relies on projecting simple geometry in front of the user, it is not the case with the stereographic projection. For exploring virtual environments in focus+context, a fully virtual environment would naturally best. As the projection was initially intended to be viewed from a (virtual) first-person view of a city, augmented reality could bring this kind focus+context to the real world, albeit with the constraints of physical navigation.

The above can be summed up in the final hypothesis of this chapter:

> *The inverse stereographic projection with its variations is a powerful tool for exploration of potentially multilayered focus+context in diverse geospatial applications, which can fully exploit the benefits of immersive environments while being intuitive to understand.*

So far, these ideas and hypotheses are all theoretical considerations, using simplified diagrams at most. In the next part, the details of how to actually implement them into practical applications are considered: mathematics, software, hardware, interaction, data. The constructed tools can then be further analyzed, tested, and discussed.

Part III

Implementation and Evaluation

6 Virtually Physical Data Layer Stack

To test out the hypotheses made in the previous chapters, these concepts have to implemented in software applications that can run smoothly, preferably on readily accessible computers and interface with current immersive technologies, to be evaluated with users. This software should be able to display sufficiently large areas of urban environments in fine enough detail to make an evaluation of the hypothesized methods' suitability for multilayered and focus+context visualizations possible. Preferably this data should not be baked into the system, but rather be accessible on demand to enable evaluations with arbitrary regions.

This chapter starts with describing the choices made regarding hardware and software for the implementation of the MCV stack, and how it interfaces with existing solutions for storing and accessing geospatial information. This is followed by evaluating the implemented system in a formal experiment, the results of which are then discussed to form a broader conclusion about visualizing multilayered geospatial urban data in virtual reality.

6.1 Implementation of the MCV Stack

6.1.1 Tools and Data

3D video game engines are finding increasing use for constructing visualizations [Buy+17] and simulations [Cri+15] in urbanism as well as other fields, and are particularly well suited for developing applications for immersive analytics — researchers have already created entire toolkits for that domain [Sic+19; Cor+19]. They provide a full development environment ready to import, generate, create, and animate 3D (and 2D) environments by exposing shaders and graphics libraries with fully-featured APIs to a variety of modern programming languages. This can significantly speed up development time, without necessarily impacting performance, as they are flexible enough to allow programmers to write their own shaders and use native code if necessary.

The video game industry at large being split between using Epic Games' *Unreal Engine* [Epi20] and Unity Technologies' *Unity 3D* [Uni20], there exists a clear preference among

smaller developers, hobbyists, and (most importantly for this case) researchers for Unity 3D; e g. all work cited in the above paragraph relies on it. Some of the factors playing into this may be the far larger "community" of Unity developers, and the resulting overabundance of support forums and tutorials, estimated to be around ten times larger than that of Unreal.

Both Unity and Unreal support development for most current XR devices. From this growing field of hardware, the industry-standard first-generation HTC Vive VR setup was chosen for its reliable performance, excellent controllers with accurate six-axis tracking and precise analog triggers and multifunctional touchpads (Figure 6.1), and for its compatibility with highly accurate eye-tracking devices. Its maximum resolution of 1080 × 1200 pixels per eye can easily be driven at its maximum framerate of 90 Hz with consumer-grade graphics cards (the NVidia GeForce RTX 2070 in this case). The choice of VR over AR for this application comes down to the much larger available field of view even in this first-generation VR headset over that of Microsoft's latest Hololens 2 (110°over 52°diagonal, respectively), and the complete control over the environment that VR provides, which can create a large and clutter-free background to place virtual objects such as the MCV layer stack.

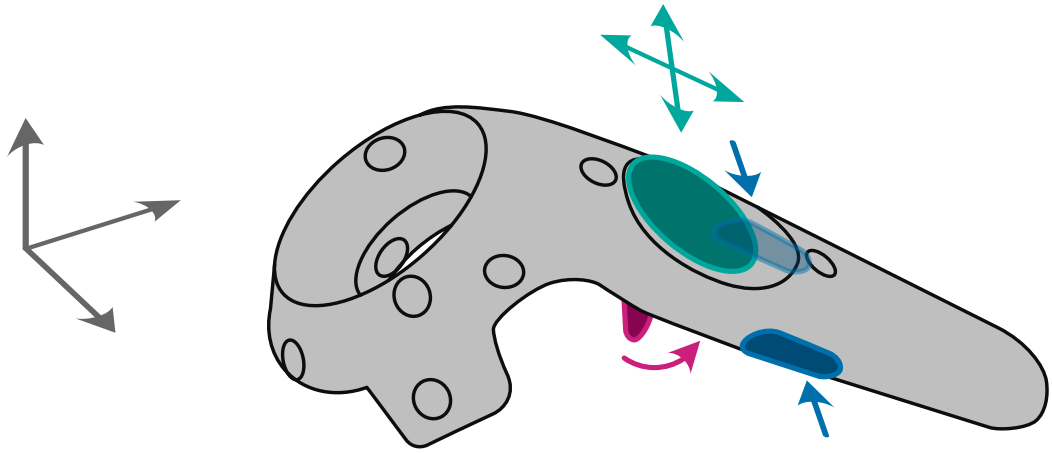


Fig. 6.1: One of the two (identical) controllers of the HTC Vive set. It is tracked in translation and rotation, and has three main input methods: the trigger (magenta), the side or "grip" buttons (blue), and the touchpad (green). The trigger outputs its continuous depression level, and the grip buttons output a binary value if any side is depressed. The touchpad registers the x–y position of a finger on it and can output a binary state when it is clicked.

Where the two game engines differ however is their potential integration with geospatial data, as only Unity 3D is supported by a major mapping service provider with a fully-featured SDK. That provider is *Mapbox* [Map20], which can be used as a platform not only to access its own cartographic data, but to author, upload, edit, and stylize custom geospatial datasets in raster and vector formats. They provide an SDK specifically for Unity 3D, which allows real-time access to any of the built-in and custom maps at different

zoom levels, and methods to convert its internal data formats (tiles and styles) into 3D objects in Unity, which can then be manipulated and animated any which way.

With these three tools — Mapbox, Unity 3D, and the HTC Vive set — it is feasible to fully implement and evaluate the MCV stack idea described in section 4.4. The data used to populate the multiple layers were taken from Mapbox [Map20] Streets v7 *tileset* for streets, river outlines, and points of interest; from OpenStreetMap [Fou20] and from the local metropolitan open data portal [Mét20] for transportation networks; and from the BD TOPO reference maps from the French National Geographic Institute (IGN) [IGN20] for building footprints, heights, and semantic information. In addition, two more data sources were used for the use case used to evaluate the system, which will be discussed in that part. All geospatial data was input into Mapbox for direct access through its API within Unity.

6.1.2 Implementation in Unity 3D

The Mapbox SDK for Unity works by requesting the specified data (vector and raster) from the server in *tiles*. A *tile provider* takes coordinates, a zoom level, and the name of the tileset, from which it creates map *objects*, which are made of a number of square planes (or meshes, if terrain modeling is requested) — as many as are needed to fill an area specified by an *extent* (the length of one side of a square in tiles). The tile provider asks for materials (which themselves consist of a shader and its configurations) and uses them to form the tiles and map textures to in the case of raster data. Vector data can be interpreted in predefined ways to create meshes, e. g. lines can be turned into polygons with a given width to create roads, and polygons with height information can be extruded into buildings that also adhere to the terrain elevation.

There are two shortcomings in how that SDK operates for the application of an interactive MCV stack. First, loading these tiles and processing them into objects for Unity can cause significant performance issues, especially when raster data is being loaded: Unity can not do texture operations asynchronously, which will cause the main thread to be blocked and result in a frozen program state until all new map tiles are successfully created. If these interruptions take too long, they will cause noticeable decreases in frame rates, up to the point of the program stalling. For the goal of using this system to evaluate the concept of the MCV stack, these issues would be unacceptable, as such perceived "jerkiness" will significantly degrade the quality of experience and thus distract from other aspects of an interactive system [SSM13].

This problem can be circumvented by pre-loading all tiles at a sufficiently detailed zoom level that will be necessary for a specific use-case or scenario at launch of the software, which will accordingly take longer, but eliminate performance issues during actual use.

Each new scenario would of course then also require a period of loading any new tiles that are required for it.

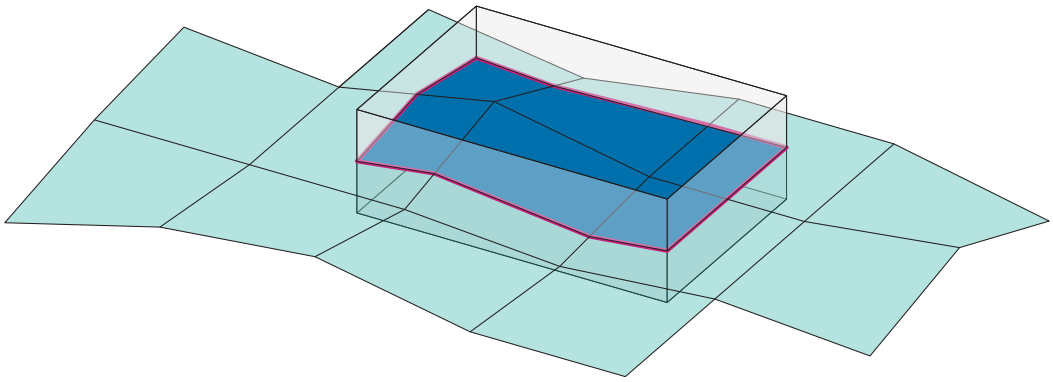


Fig. 6.2: A bounding box (grey), which defines a frame within which map tiles (green) are displayed (blue). The user will only see the part of the map that intersects this box.

This solution however plays into the second shortcoming of the tile provider: there exists no functionality of "framing" the created tiles; all loaded (and enabled) tiles are shown in their full extent. To fit the layout of the MCV stack as shown in Figure 4.9, or in fact any other MCV system, each view must be clearly framed, so as to not interfere with any other, and not to needlessly fluctuate in size. The solution chosen for this system is to create a shader that culls all geometry from a specified bounding box from projection, and assign it to the map tiles. This creates a three-dimensional frame with user-defined dimensions (Figure 6.2), within which the tiles can be freely moved around for panning the map, without occluding other parts of the scene, and without changing its shape, as long as the movement is constrained in a way such that there are always tiles inside the frame.

Map navigation was implemented using one of the HTC Vive controllers in several ways. The first is an indirect "grabbing" method: when holding the grip button on the controller (Figure 6.1), its movement in 3D space is converted to movement of the map object. Left–right and forward–backward motions are result in panning: the map translates horizontally relative to its frame, following the controller. Zooming is accomplished by moving the controller up and down while holding the grip button, which scales the map around the center of its frame. This way, the "grabbing" method can freely pan and zoom in a single motion that should feel natural. Very similar indirect navigation methods have been successfully demonstrated and evaluated for wall-sized [Nan+11] and hand-held displays [Jon+12] as well as desktops [SK17], and more recently for immersive applications [Sat+19], where they were shown to reduce fatigue over direct interactions that require touching the display or map itself.

A more "direct" method of panning is to use the trigger: when slightly depressing it, a visible ray is emanated from the tip of the controller, that can intersect, or "hit" the visible part of a map inside its frame. When the trigger is fully depressed while hitting a map, it

can be dragged along horizontally following the controller's motion until the trigger is released again. Zooming can also be more directed this way, by swiping up and down the touchpad while the ray hits a map, which then gets scaled around that hit point. Swiping the touchpad without engaging the ray results in centered zooming, as with the grabbing motion.

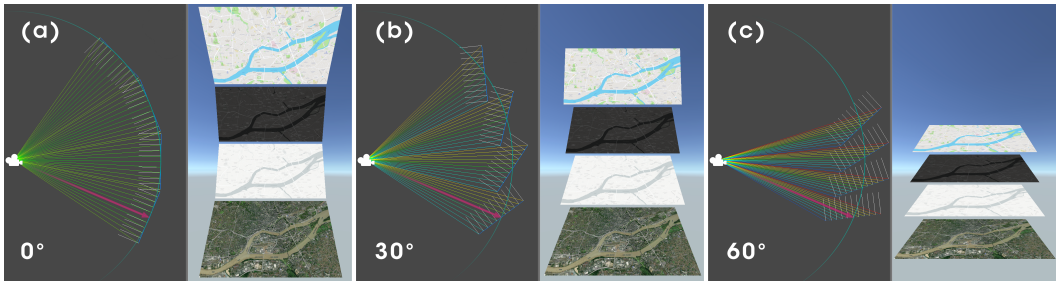


Fig. 6.3: Different viewing angles demonstrated for the MCV stack. The center of each layer is at the same distance to the viewer, the arc between layers is the same. Rays from the camera to the layers are colored by their length relative to the central ray. White lines show the normals the viewing angles are referenced to.

To create the MCV stack, multiple map frames as described above can be placed in the virtual space according to the scheme described in section 4.4 and shown in Figure 4.9. The effects of navigation methods are mirrored across all layers: the offsets and scaling applied by the "grabbing" method are simply copied to every layer. Whichever layer is hit by the ray for the "dragging" method becomes the "master" layer, from which the navigation is then copied to the remaining layers.

To ensure an ideal placement of the layers, in which every layer is presented equally to the viewer, the layout must be updated automatically in adherence with the following parameters:

- The viewing *angle* in the center is the same,
- the apparent distance *between* each layer is the same, and
- the *distance* from the center of each layer to the viewer is the same.

This is demonstrated in Figure 6.3, where alterations in the viewing angle change the layout to adhere to all parameters, although distortion from a wide-angle projection of the layout makes the layers appear uneven. The algorithm as implemented takes input from the user regarding the dimensions of the frames, the viewing angle, the arc of the distance between layers, and the height of the lowest layer from the floor. The distance is taken from the lowest layer's distance to the viewer, which can change with every movement of the viewer. Every layer's orientation is adjusted accordingly, along with the position of each layer above the lowest one. Alternatively, instead of keeping the lowest layer at a fixed position, it can follow the viewer horizontally at a set distance, so that only vertical movement affects the layout.

This implementation also works in a desktop environment, where a mouse or touchpad can be used to emulate the functionality of the Vive controller. The focus of this work is however on immersive environments, not least because their evaluations are currently being considered "grand challenges" in the field of immersive analytics [Ens+21]. A stack of layers for map comparison has not yet been formally evaluated according to literature review, much less one implemented for immersive environments.

6.2 Experiment: Comparing MCVs in VR

In the motivations for the MCV stack (section 4.4), it was contrasted against another juxtaposition method — a regular grid of layers — with the hypothesis being that the stack would compare favorably against it, especially with higher numbers of layers, being more compact and more efficient at using the virtual space, while also showing a more coherent view of the different layers. Another map comparison technique (section 4.3) would also struggle with larger numbers: temporal multiplexing, or *blitting*. Unless e. g. small multiples are employed as previews of the different available layers, a user's memory and attention would be taxed quite highly to keep track of the different layers and make comparisons between them.

With low numbers of layers, these three techniques could however become more comparable. A grid of e. g. four layers would not take up much more visual space than a stack of four equally-sized layers, just extend more to the sides than vertically. A blitting interface could use the same space for a larger display of its single layer, while switching between just four layers would not quite reach cumbersome levels, especially with a fast interaction method that lets the user switch quickly between any two layers.

If these three techniques are implemented equally well in the same immersive environment, comparisons between them could come quite close, perhaps even down to personal preference. Even though the stack retains its theoretical advantages of a more coherent and compact layout with the added functionality of a connecting line between layers, the other two techniques could prevail in a "fair" evaluation due to their more familiar nature, or the larger view of single layers.

In Schulz' *A Design Space of Visualization Tasks* [Sch+13], the concepts of data, tasks, and visualization are combined in two ways to ask different questions:

- *Data*+ *Task* = **Visualization?** and
- *Data*+ *Visualization* = **Task?**

The first combination asks which visualization needs to be created for a given task and data, whereas the second can follow as an evaluation process, once a visualization has been defined: how well can tasks be performed on that data using this visualization?

To perform this evaluation, the task and data have to be well defined. Usually, the effectiveness of geovisualization systems is evaluated with simplified or synthetic tasks, such as finding certain features on one map, or differences between two [LPA15]. While these methods can often be generalized to map legibility, the aim of this work is to more directly test how well a system can facilitate an understanding of multiple layers of spatial data. To do this, a task for a user study that can measure such understanding, or at least ask the user to embark on the process of understanding the interplay between different layers before doing an action needs to be devised. Preferably, such a task would also be close to real-life studies in urbanism, and ideally even contribute to them, in addition to geovisualization research. The following section describes how such a task was found in the use case of the public perception and participatory improvement of street lighting.

6.2.1 Use Case in Urbanism: Public Street Lighting

The practice of artificial lightning is based on technical and political choices developed over several centuries. Since the 1970s, it has been regularly questioned, first by observing the worsening of light nuisances which disrupt astronomical work, then by the increase in energy consumption due to the expansion of the public lighting fleet. In an unprecedented situation, partial or even total extinguishing of public lighting is now under consideration for towns. Some small French municipalities have opted for this practice, which remains relatively complex to imagine in the dense urban fabric of a metropolis, particularly because of the intensity of nocturnal work shifts and personal practices. Where public lighting has been built in a systematic way, its evolution is now envisaged by urbanists through empirical approaches based on the acquisition of perceptive data from users of the public space at night.

For a community, public lighting is a tool used to meet several objectives: securing goods and people, ensuring comfortable mobility, promoting architectural heritage and landscapes, and accompanying nocturnal urban activities. In the current model of conceiving urban ambiances as used by teams of decision makers, planners, and installers, the actual *users* of a public lighting service, who evolve within the nightly ambiances that were built for them, have had no influence so far.

New considerations of energy factors have led communities to greatly modify their approach to artificial urban lighting — it must now meet the traditional criteria regarding security, mobility, and identity, while also committing to energy-saving, environment-friendly, and healthy strategies. This topic, which has recently been the main subject of one of the *Great Debates* initiated as a measure of public discourse by municipal governments [Gam17], seems to have become a general concern. The consequences of increasing illumination are being studied, with their impact on fauna and flora [KH13] and on human health [Zie18] being well documented.

The usually technical operation of public lighting thus finds itself facing more sensory approaches. Its protocols in terms of political and technical decisions must now conform to questions of coexistence between species or of preservation of the night sky [Cha11], while also meeting recent regulatory requirements. The technical expertise thus encounters an increasing number of contradictions and balancing acts, e. g. illuminating a mobility space located at a water edge, staging the built and natural heritage, and respecting the darkness needs of ecosystems.

Here, citizens could provide arbitration elements from their experiences, being experts in their own uses of these territories at night. The problem however remains these users' knowledge about those contradictions. To potentially rank the political and technical choices associated with lighting, the users of public spaces need to be informed clearly about various criteria to meet, for which an efficient way to provide that information is needed, as well one to solicit their feedback in order to contribute to improving public street lighting.

A part of this rather complex task could be conducted as part of a formal study involving the layer stack and comparable MCV methods, where relevant geospatial data is presented in the different layers to users, who are asked to provide their feedback after studying these layers. This feedback could come in the form of identifying "problematic" areas in terms of public lighting after being made aware of the above issues, and a written explanation of this identification. Splitting information across layers in MCVs instead of attempting to merge all into one map also allows to assess each user's attention to the individual layers by employing eye tracking technology. With this use case and the three different MCV methods available to provide the visualization necessary for feedback solicitation, a formal study can be designed to aim at multiple simultaneous goals:

1. Determining the differences in how users perceived each MCV through questionnaires after tasks and analyzing eye tracking data,
2. seeing if these immersive MCVs are being useful in an actual participatory study with questions regarding urban design, and if the choice of MCV had an impact on the quality of the answers participants provided, and
3. the goals of the experimental urban design study itself:
    a) Assessing the knowledge and competence of participants regarding public lighting,
    b) asking them to decide on which parts of the city could benefit from reducing public illumination, and
    c) the importance participants give to the topics of mobility, of respect for the night sky, of energy conservation, and of easy access to nocturnal activities when making these choices.

To present the relevant information for participants to give their feedback, four layers of geospatial data are necessary, showing:

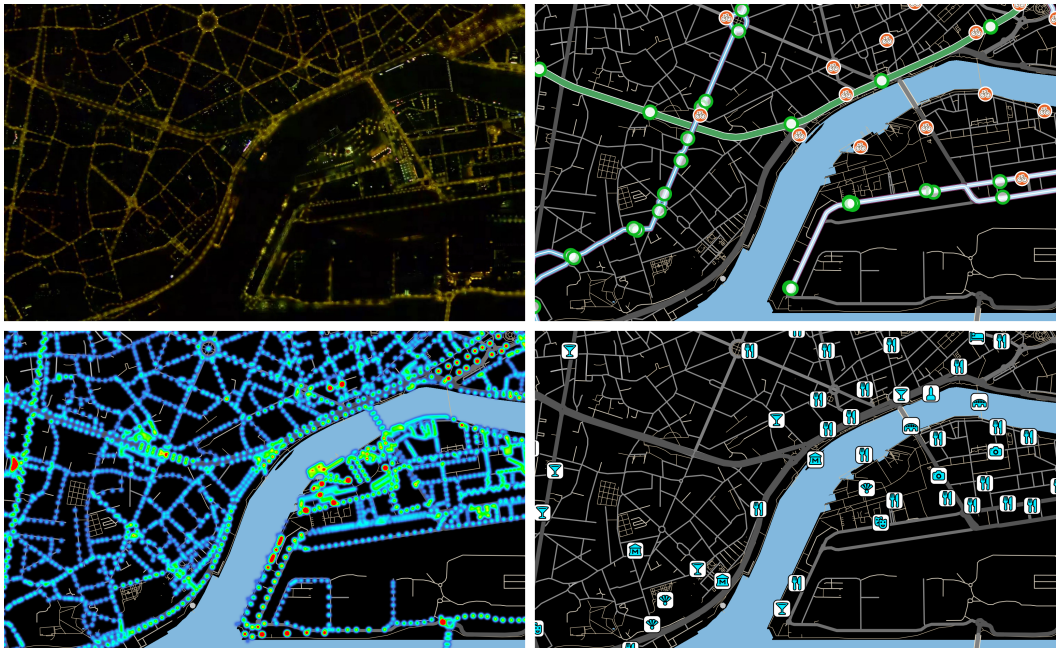


Fig. 6.4: Clockwise from top left: light pollution aerial orthophoto, nightly transportation network, nocturnal points of interest, and a heatmap of energy consumption by public lighting.

- the energy consumption (EC) of each individual street lamp, visualized as a heat map ranging from blue to red,
- the light pollution (LP) caused by artificial light sources, seen in the form of aerial photographs taken at night,
- the transportation lines and stops that are active at night time (NT), along with bike-sharing stations; with colors and shapes matching those of the local transit maps, and
- points of interest (POI) operating at night, such as restaurants and bars.

These data layers (shown in Figure 6.4) are particularly well suited to an MCV approach instead of superimposing them into one view, as most of their information lies along streets. Superimposition here would lead to severe crowding, visual clutter, and occlusion.

6.2.2 Study and Task Design

Gleicher argues that "much, if not most of analysis can be viewed as comparisons" [Gle18]. He describes comparisons in the broadest sense as consisting of items or targets, and actions — what a user does with the relationship between items. These actions have in earlier literature on comparison been limited to identifying relationships, but Gleicher extends the categorization of comparative tasks to include measuring/quantifying/summarizing, dissecting, and connecting relationships between items, as well as contextualizing and communicating/illuminating them.

In the use case, one can see how a user would do most of the above actions comparing the data layers—the items/targets in this case—to come a conclusion and communicate it. Following Gleicher's considerations on what makes a "difficult comparison," the task needs to be refined under the following criteria:

- the number of items being compared,
- the size or complexity of the individual items, and
- the size or complexity of their relationships.

The first two issues were directly addressed by simplifying the choice a user had to make. The area of the city for which data coverage in all four layers was available was divided into twelve similarly-sized regions—one for each *scenario*. In each, three items were outlined: the *candidate* areas, out of which a user would then select the most "problematic" one — the one most in need of reduction of public lighting, or an overhaul in efficiency. This considerably reduces the number of items, their size, and complexity, in comparison to e. g. letting users search larger parts of the city and creating their own outlines. Having a fixed number of candidates in all scenarios allows for consistent comparisons of completion time other "performance" metrics, such as the number of times participants switched their attention from one candidate to another.

To address the third point (the complexity of relationships), special care was taken to make each set of three candidates as comparable as possible from a geometric point of view — three roundabouts, or three crossroads, etc. — to draw attention to their differences in their data layers, and lessen the impact of their physical characteristics. An example: one scenario's candidates are all road segments that include one roundabout, and have similar energy consumption. One of these is close to a tram stop, and another has stronger light pollution. The aim to reduce the need for users to balance too many aspects at once, while providing insights during the urbanist analysis regarding the remaining differences that mattered most in the participants' decisions.

6.2.3 Experiment Setup

In order to turn this task into a controlled user study, a careful setup of the experiment was necessary, in terms of the software as well as its procedure. First, the three systems — the MCV stack, the grid, and the blitting system — were arranged to be comfortable to view and use. The configurations were not mathematically designed to make each system "equivalent," as the different distortion characteristics (large vs. small layers, tilted vs. flat, etc.), and the differences between seeing only one or all four layers at the same time are rather difficult to quantify and balance. Instead, the configurations were iterated on with a group of urbanists during the development phase and in pre-tests, asking for their feedback and adjusting the systems in real time with the goal of making the systems appear *subjectively* equivalent in their usefulness of visualizing four layers:

each system's configuration was tuned to maximize the experts' satisfaction with them. Another argument against a more mathematical approach was the expectation that users would each take slightly different positions and move around, which would render more precise optimizations less effective. The goal was to make viewing and interacting with all three systems as comfortable and equal as possible, given their differences, in order to allow a fair comparison between them.

From this process, the following decisions regarding the configurations of the MCVs were made:

- The MCVs would be placed within a semi-realistic virtual room, instead of floating in a void, to provide a modicum of environmental familiarity and prevent a sense vertigo, which was observed in some users.
- The MCVs would occupy a fixed position within that room, and not follow the user around to keep a fixed distance, as users tended to leaning in for closer inspection, and movement of the layers was seen as unexpected and unnatural.
- For comfort during a rather long study, the participants were to sit rather than stand.
- All layers were fixed to a 3:2 aspect ratio.
- The blitting system's single view measured 1.5 m × 1 m, approximately 2 m away from the user, and was oriented perfectly vertical.
- The grid system's four views measured 1.125 m × 0.75 m each, approx. 2.2 m away, and tilted by 10 degrees.
- The stack system's layers measured 1.2 m × 0.8 m, at a distance of approx. 1.8 m.
- All views hover slightly (approx. 20–40 cm) above the floor.
- The navigation was simplified: only the "indirect grabbing" scheme was kept from section 6.1.2, which would also remove another variable for a more direct comparison.

The second point also led to a change from the original MCV stack implementation as discussed in section 6.1.2. Instead of constantly adjusting the positioning of each layer to always face the viewer at the same angle and at the same distance as shown in Figure 6.3, its configuration is static and optimized for the default viewing position of a seated user. This change also makes the stack more fairly comparable to the other two systems, which are static as well.

The last point frees up the input methods on the HTC Vive controller to include more functionality. As described in section 6.1.2, the side buttons are used for the grabbing method to navigate maps. This method has been slightly modified, in that the controller's motion being either parallel or perpendicular to a map now dictates whether it pans or zooms. In the case of the stack, the lowest layer is used as a reference for this.

The trigger no longer being used for the dragging method, it now is used to project the same ray, but only to place a small, spherical marker on one layer, which will also be replicated across the other layers. In the case of the stack, those markers are also connected vertically with light rays. Lastly, the touchpad is now used to switch between the different layers in the blitting system, with an annular menu hovering just above it that contains symbols representing each layer; the same symbols are also present in each system next to the according layer to reinforce their connection. Swiping across the pad with a thumb lets the user quickly switch directly from one layer to any other. Figure 6.5 shows how these three systems appear to a user in VR.

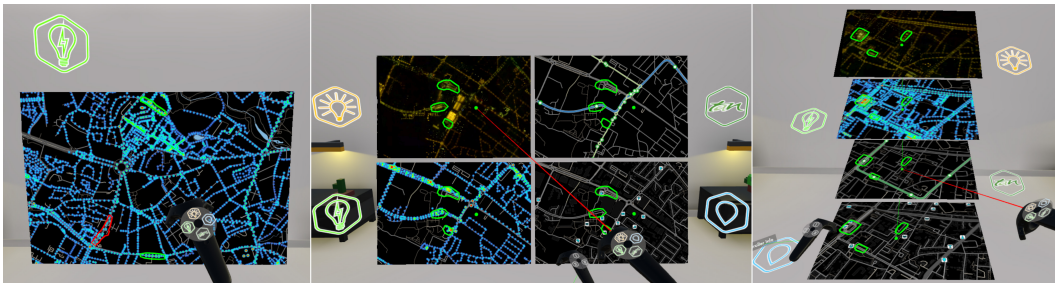


Fig. 6.5: The configurations for the blitting (left), grid (middle), and stack (right) MCV systems as used for the user study. Marker placement is shown for the grid and stack — note the connecting line in the latter's case.

This way, one controller is now fully dedicated to navigating across and between layers of each MCV. The second controller can thus function for the other part of the experiment: navigating and selecting candidate areas. It was setup to use its touchpad with an annular menu similar to the first controller, but to switch the highlight between different candidates. Each candidate is assigned a number 1–3, in a clockwise arrangement. The annular menu contains these numbers, and touching it changes the corresponding candidate's outline from green to red, while also changing the number floating above the map.

The second controller's trigger can be used to continuously fade the outlines of all candidates in opacity to reveal any information occluded by them. Once the user wishes to finalize a choice for a candidate, the touchpad needs to be pressed down on that number for a moment, until a confirmation interface appears: a red cube, offset by 20 cm away from the controller's last position appears. To confirm the tip of the controller has then to be moved inside this cube and held there for three seconds, until it turns green. This process is shown in Figure 6.6.

Finally, the side buttons of the second controller are used for advancing through different steps in the experiment procedure as discussed below, and also for showing information about the controllers' functionality via hovering tooltips, as well as showing additional information about each layer.



Fig. 6.6: The interface for selecting and confirming candidate areas.

At the beginning of each session, a text field floats above the controller used for map navigation, instructing users to take it with their dominant hand, as the fine movements related to navigation and pointing require more dexterity than the actions of selecting and confirming candidates.

6.2.4 Experiment Process and Participants

The study, following a within-subject design which exposed each participant equally to all three MCV systems, was conducted in 26 sessions with as many participants (9 female, 17 male), with ages ranging from 18 to 45 years ($M = 21, SD = 6.33$). Figure 6.7 and Figure 6.8 show their demographics, habits, and experiences regarding city and map navigation, 3D video games, and virtual reality. Two particular participants have been highlighted, as they will be discussed in the next sections as examples of the analysis of results from an urbanist's point of view to answer goal 3 listed in section 6.2.1.

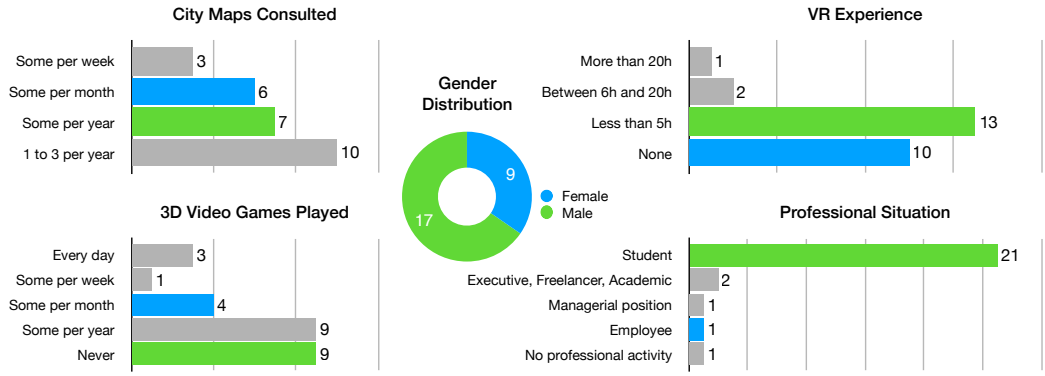


Fig. 6.7: Histograms showing the demographics and habits of the 26 participants. Participant 1 and 21 highlighted in green and blue, respectively.

As explained in section 6.2.2, the city was divided into twelve different parts, resulting in twelve *scenarios* for each session, each containing three outlines demarcating the candidate areas. Pairing one scenario with one MCV system resulted in one *stimulus*. One specific stimulus was chosen to always be used for the first tutorial to introduce the data layers and explain the interaction methods on the blitting system, based on its characteristics. It showed the most well-known part of the city, and seeing all candidate zones required panning the map with the controllers. The other eleven scenarios appeared



Fig. 6.8: Histograms showing the demographics and experiences (measured by visual analog scale, 0–100) of the 26 participants. Participant 1 and 21 highlighted in green and blue, respectively.

in a random order for the remainder of each session: two for the remaining part of the tutorial in order to introduce the other two systems, and nine for the actual tasks which were used to evaluate the system. This splits each session into a tutorial part, and an evaluation part.

Each scenario itself is further divided into an instruction phase, a task phase, and a questionnaire phase, in which the participants reflect on the just completed task. The instructions consisted of interactive, step-by-step tutorials explaining each function of the controllers, the meaning of each layer, and how to navigate them during the tutorial part, and quick reminders during the data-gathering part.

Before the tutorial part, each session was started with an introduction to the study, in the form of two pages given to read (their full text is in the annex at section 10.2), preceded by a visual acuity test and a screening for color blindness and followed by a first questionnaire concerning demographics and experience/habits regarding cities, maps, video games, and VR (shown in Figure 6.7 and Figure 6.8. The participants were personally introduced to how to fill out the digital questionnaires, which were running on a separate PC, and how handle the VR headsets and controllers, from which point on everything else was autonomous.

As said before, the tutorial begins with the same scenario and the blitting system for each user, to ensure maximum consistency in their training. The blitting system was chosen here for being likely the system that most closely matches what participants are already familiar with from using digital maps, and because it only shows a single layer at a time it is best suited to explain each layer without interference.

At first, all of the controller navigation functions are disabled. The tutorial system then gradually introduces and enables panning, zooming, and blitting the map layers, and selecting and confirming candidates by asking the participant to perform simple tasks and waiting for successful completions. While introducing the blitting mechanism, each layer was explained in detail; participants could not switch to the next one before confirming their understanding of the description. The first such training scenario then introduced the task of selecting and confirming one of the three candidate zones, followed by the instruction to take off the HMD and processed with the questionnaire, which itself introduced the types of questions that were to be answered throughout the session.

The second and third scenarios were used to introduce the grid and stack systems in a similar fashion, only forgoing the complete explanations of each layer. They similarly ended with the actual task and the questionnaires. Here, the order of the systems was balanced between the participants: half were first exposed to the grid after the blitting system, and the other half to the stack. The map regions were randomized from the second scenario on.

The remaining nine scenarios were divided into three blocks of three stimuli: a repeating permutation of the three layering systems ((B)lit, (G)rid, (S)tack), balanced across the participant pool; e. g. GSB-GSB-GSB, or BGS-BGS-BGS. The tutorial system was pared down to only instruct participants in the blitting system to switch between all layers at least once at the beginning, and to show reminders of the controller functions, as well as the layer descriptions if requested. Instructions appeared to evaluate the scenario, pick a candidate, and proceed to the questionnaire. While a participant was away from the VR headset and busy with a questionnaire between tasks, the system would automatically load the next scenario and set up the next introduction, with reminders about the task.

This choice of procedure — always introducing the blitting system first — could appear to impact the balancing of the study. However, presenting that system first made it possible to create an easier in-system tutorial for introducing each layer separately, as opposed to possibly overwhelming the participants with all data layers at once. In this aspect, it is also the simplest system to start with for explaining the interaction methods necessary to navigate the maps and selecting candidate zones. As stated above, it is also most likely the closest to previous experiences participants might have had with multilayered maps: spatial multiplexing is rather uncommon outside of professional applications, whereas typical digital maps usually offer an interface for switching between data layers. The

participants are thus assumed to already be "unbalanced" in this regard: the blitting system does not introduce any new concepts to them other than the navigation in VR, which is shared among all systems under evaluation. Therefore, this approach of introducing the blitting system first, and then randomly the grid or the stack second and last, appears to be in fact the most "balanced" one possible, considering real-life experience. The two "novel" systems were introduced in a true balanced way, and the evaluation part after this tutorial part was traditionally balanced using a Graeco-Latin square design.

6.2.5 Devices, Measurements, and Questionnaires

Two separate devices were used for the study: one PC handling the execution of the MCV systems through the HTC Vive VR setup, and a second one with the sole task of hosting the digital questionnaire, which was created with a tool developed by Guse et al. [Gus+19]. This questionnaire was also used to instruct participants in how to proceed with their tasks. This way, all interactions for the complete duration of the evaluative sessions were fully automated and as equal as possible between participants: the questionnaire instructed participants to put on the headset, the MCV system told them when to take it off again and go back to the questionnaire, and so on until all tasks have been completed. Switching between VR and a simple questionnaire on a regular desktop PC screen provided the benefit of giving the participants regular pauses from wearing the headset, the excessive use of which might have lead to needless additional fatigue.

The first-generation Vive headset used for the study was equipped with the SMI binocular eye-tracking retrofit, capable of sampling gaze positions at 250 Hz with a reported accuracy of 0.2 degrees. The system ran on a Windows workstation powerful enough to render all MCVs at the Vive's maximum frame rate of 90 Hz and its maximum resolution of 1080×1200 pixels per eye at constant performance.

The interactive questionnaire ran on a separate PC inside a Chrome browser window, close enough so the participants could stay in their swivel chair to access it without standing up, but far enough not to interfere with the VR tasks. Aside from gathering demographic information at the beginning, it recorded subjective data by asking the participants about their reasons for choosing the candidate areas they did after each trial, as well as by providing text fields for free-form feedback about each individual system after all scenarios were finished. In this post-hoc phase, it also presented pairwise preference questions about the three MCV systems under three aspects: legibility of the map layers within these systems, their ease of use, and their visual design.

To complement this declarative feedback, the VR system also recorded performance aspects (completion times and interactions) and visuo-motor data via oculometry and tracking of the headset. Eye movements bring a wealth of information—they are overt clues about an observer's attention deployment and cognitive processes in general, and are increasingly being tracked for evaluating visual analytics [Kur+16]. In the context of map reading, measuring gaze allows to find out precisely which map layers participants chose to observe in particular, and at which times. Furthermore, gaze features and their properties, such as saccades and fixations can be derived, in this case by processing with the toolkit developed for the *Salient360!* dataset by David et al. [Dav+18], and further refined using a velocity-based parsing algorithm [Dav+19; SG00].

6.3 Results

6.3.1 Preferences and Rankings

The participants were asked about their opinions regarding different aspects of the three layering systems as some of the last questions of each session. To quantify their preferences, pairwise comparison questions (randomized and balanced in their order) were used to create pairwise comparison matrices (PCMs) for each of these three aspects:

1. **Map legibility:** which system showed the map layers in the clearer way and made them easier to understand for you?
2. **Ease of use:** which system made interacting with the maps and candidate areas easier for you?
3. **Visual design:** which system looked more appealing to you?

The resulting PCMs could then be used in different ways. First, rankings were determined for the three MCV systems by counting the number of times one "won" against the others in the comparisons under each of the three aspects. In that system, MCVs "winning" twice were called the "preferred" MCV system by the participant under that aspect. Winning once placed them second, and zero wins resulted in it being placed last, or what was called the "disliked" system. These results are shown in Figure 6.9 a.

By this measurement, more than half of the participants placed the blitting system *last* under each aspect. The grid was almost evenly split between being rated first or second under the *ease of use* and *readability* aspects, while being a clear second in terms of visual design, where the stack system proved to be the favorite. The stack also shows a slight advantage with its number of *firsts* in terms of ease of use over the grid.

These PCMs were also fed into a Bradley-Terry model [Bra84] to generate relative scores for each system, shown in Figure 6.9 b. These results mostly mirror the results from the

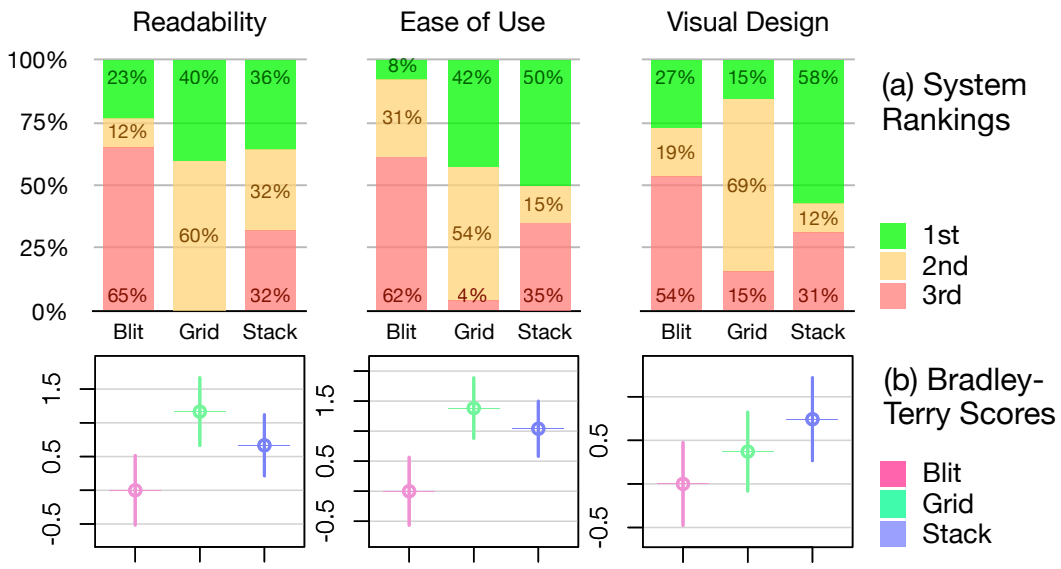


Fig. 6.9: **(a)** How often (in percent) each MCV system was ranked *first, second,* or *last* in terms of readability (legibility of the map layers), ease of use, and visual design by the users; **(b)** Pairwise comparisons scored with the Bradley-Terry model.

first ranking approach in Figure 6.9 (a) in a less granular way, but with a key difference: it gives the grid system a clearer advantage over the stack in terms of readability and, to a smaller degree, ease of use. This can be explained by the evidently polarizing nature of the stack, which stands in stark contrast to the grid. The grid was rarely if ever rated last in any aspect, and usually attained second place, whereas the stack garnered the highest number of first place ratings, as well as a considerable number of last places. This "love it or hate it" kind of response can also be seen in the blit system to a lesser degree, in that it similarly had a very low number of second place ratings, but it was also rated as being worst far more often than being first.

Condensing those pairwise comparisons further, it was counted how many times the participants ranked each system first and last across the three aspects; results shown in Figure 6.10 a. This method shows how the blitting system received zero *first place* rankings from the majority of the participants—under no aspect was it the *best* among the three systems for them. It also shows how only the stack system ever received a *first place* under all three aspects—in fact, from a quarter of the participants.

The chart showing the number of *lasts* received allows similar observations: the blitting system received the maximum number of *lasts* most often, and the stack had an overall lower number of *lasts* than it. The grid system however received by far the least number of *lasts:* only a fifth of the participants deemed it to be worse than both other systems under a single aspect. Most participants thus found either the blitting system, or less frequently, the stack system to be worse than the grid under at least one aspect; once again showing it to be a "middle ground" kind of MCV system.

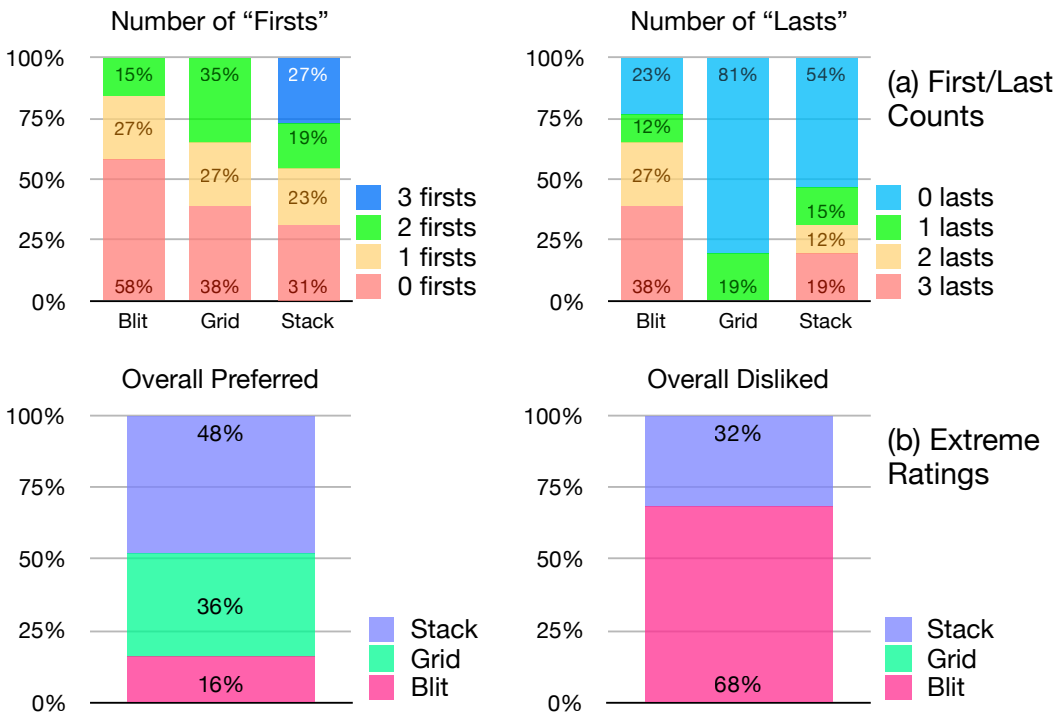


Fig. 6.10: **(a)** Which proportion of participants gave the MCV systems a number of zero to three *first* or *last* rankings across all aspects (readability, ease of use, and visual design); **(b)** Which proportion gave the most *firsts* (and therefore *preferred*) or *lasts* (and therefore *disliked*) to each system in the pairwise comparisons.

In a last distillation of the PCMs, it was measured how many of the of the total extreme ratings (*firsts* or *lasts*) as per the previous measure above each system received, allowing a presentation of the overall *most preferred* or *most disliked* systems in Figure 6.10 b. By this measure, the stack MCV was preferred overall by almost half of the participants, followed by the grid system with more than a third. Two thirds of the participants showed a strong dislike for the blitting system, and none of them did so for the grid; once again demonstrating its noncontroversial nature.

From the demographics gathered in Figure 6.7 and Figure 6.8, the most interesting characteristics of participants appeared to be frequency with which they played 3D video games and consulted city maps, in that these could be well juxtaposed with their system preferences; shown in Figure 6.11. Out of those identifying as never playing such games, most prefer the grid system, and some the stack—none picked the blitting system as their favorite as measured by the method in Figure 6.10. The MCV stack shows a tendency to be favored by those who play at least a few times per month. Those who use city maps most frequently (at least multiple times per month) disproportionally prefer one of the spatial multiplexing systems, especially those who most rarely use maps preferred the stack system most often. Concerning the *disliked* system charts in Figure 6.11, the blitting system overwhelmingly earned the least favorite status from all kinds of participants, however mostly from those with the least gaming experience, and

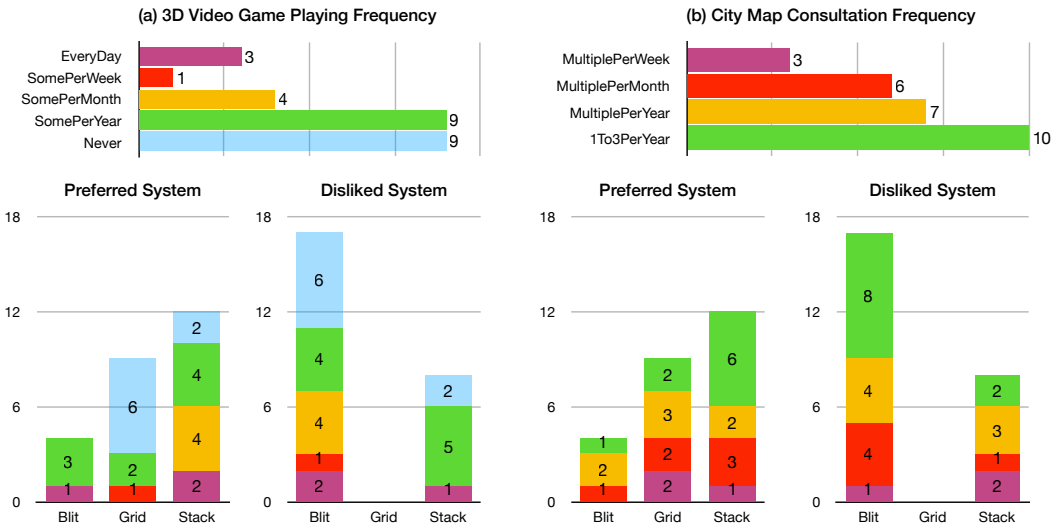


Fig. 6.11: **Left:** the number of participants who play 3D-based video games at different frequencies, and their distribution among those who *preferred* and *disliked* each system, according to the measurement shown in Figure 6.10;
Right: the same for the frequency with which participants consult city maps.

from those who very infrequently consult city maps. None place the grid system last, as seen before in Figure 6.11 b.

6.3.2 Subgrouping Results

One of the main goals (from section 6.2.1) was to explore how the three MCVs were perceived. Most important to that question is the aspect of how legible the maps and layers are when using one of these systems, which was directly asked during the pairwise comparisons. These rankings regarding this "readability" were already presented in Figure 6.9 and discussed above, but the from that first inspection *subgroups* can be formed for further analysis. The numbers of participants who ranked each system *first* in terms of readability is roughly comparable (six preferred the grid, ten the blitting system, and nine the stack), so the total number of 26 participants was divided into three subgroups: those who (p)referred the (B)litting system, the (G)rid system, and the (S)stack.

Similarly, looking at how the user characteristics as shown in Figure 6.11 seemed to interact with their preferences for MCVs, subgroups were created based on the frequency with which the participants played 3D video games, and consult city maps. The groups *zG*, *sG*, and *fG* for video gaming frequency are:

- (z)ero (never, n = 9),
- (s)ome (some per year, n = 9), and
- (f)requent (at least some per month, n = 8) 3D video (G)aming,

and for the map consultation frequency groups *rM*, *sM*, and *fM*:

- (r)ero (1–3 per year, n = 10),
- (s)ome (> 3 per year, n = 7), and
- (f)requent (at least multiple per month, n = 9) city (M)ap consultation.

Figure 6.12, Figure 6.13, Figure 6.14, and Figure 6.16 show the results of these subgroups, as well as a (T)otal measurement for comparing each MCV system across all users. They already show how the differences between these subgroups in many measurements are much more pronounced than those between the MCV systems.

6.3.3 Performance and Behavior

The measurement of how quickly participants completed each scenario (Figure 6.12 a), how many times they shifted their gaze from one map to another during each scenario (Figure 6.12 b), and how much they panned and zoomed the maps (Figure 6.13) appeared to be the most salient indicators for investigating how participants interacted with each MCV system.

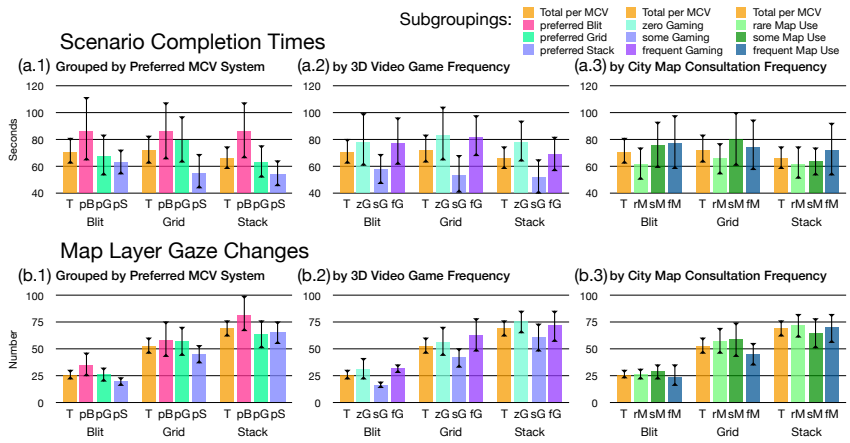


Fig. 6.12: Performance measurements for each MCV system (blit/grid/stack): scenario completion times (**a**), and the number of times gaze has shifted from one map to another per scenario (**b**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3).

When comparing the scenario completion times across the systems (Figure 6.12 a) using the (T)otal values from all participants, there are no significant differences between MCV systems. The preference subgroups here (Figure 6.12 a.1) however point toward a tendency of those who prefer the stack system (*pS*) to be significantly faster in each system than other participants, especially in the spatial multiplexing systems. Those who prefer the blitting system (*pB*) appear generally slower across all systems, albeit with a relatively high variability.

Similarly, a differentiation when grouping by video game frequency (Figure 6.12 a.2) can be seen. Significantly faster were those who play video games with a medium frequency (*sG*), while the frequent players (*fG*) were as comparatively slow as those who never play (*zG*), although with a slight tendency to be faster with the stack system. A possible interpretation here could be that those who never play felt less pressure to complete the scenarios as quickly as possible, whereas habitual gamers could have seen a quick completion as a challenge. Those playing the most may have seen the challenge not in speed, but in accuracy of their answers, or, alternatively, they could have enjoyed their time in the virtual environment more than the other groups, and not feeling the need to exit it as quickly as possible.

A very similar distribution among these two subgroups can be seen for the number of map layer gaze changes (Figure 6.12 b1, b2), but in that measurement the system itself factors significantly more. The most gaze changes happened in the stack system, and the least in the blitting system, where each "gaze change" had to be accompanied by a physical interaction with a controller. In the blitting system, the difference between the preference subgroups is quite large: those participants who prefer it change their gaze in it the most, and those who prefer the stack system do so the least. This can be interpreted in the way that those who prefer the stack rely more on their peripheral vision for comparisons, rather than directly shifting their gaze, which is easiest to do in the stack system, where the adjacent layers are much closer together, hence their preference for that one and their dislike for the other systems.

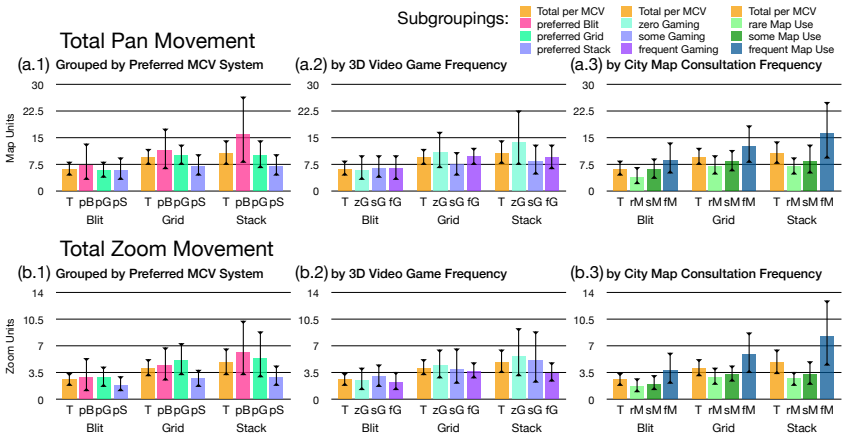


Fig. 6.13: Map navigation measurements for each MCV system (blit/grid/stack): how much they panned the map per scenario (**a**), and how much they zoomed in/out of the map (**b**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3).

The map consultation frequency subgroups (Figure 6.12 a3, b3) do not impact either measure as much as the other groupings, only showing a slight trend for those participants who consult city maps with the lowest frequency (*rM*) to spend less time in each scenario,

especially in the blitting system, and for the most frequent map readers (*fM*) to make less gaze changes in the grid system.

This subgroup is more significant in its differentiation for the pan and zoom movements (Figure 6.13 a3, b3). Generally, the participants were navigating around the maps much more in both spatial multiplexing systems than in the larger, temporal multiplexing (blitting) system, however the most frequent map users (*fM*) were clearly moving the maps much more, particularly by zooming in and out (Figure 6.13 b3)—reflecting perhaps on their experience with navigating maps. They also had a higher confidence interval in this metric, which could be attributed to a more differentiated handling of different scenarios. Otherwise, the other subgroups track with their completion times: the more time participants spend in a scenario, the more they navigate in it.

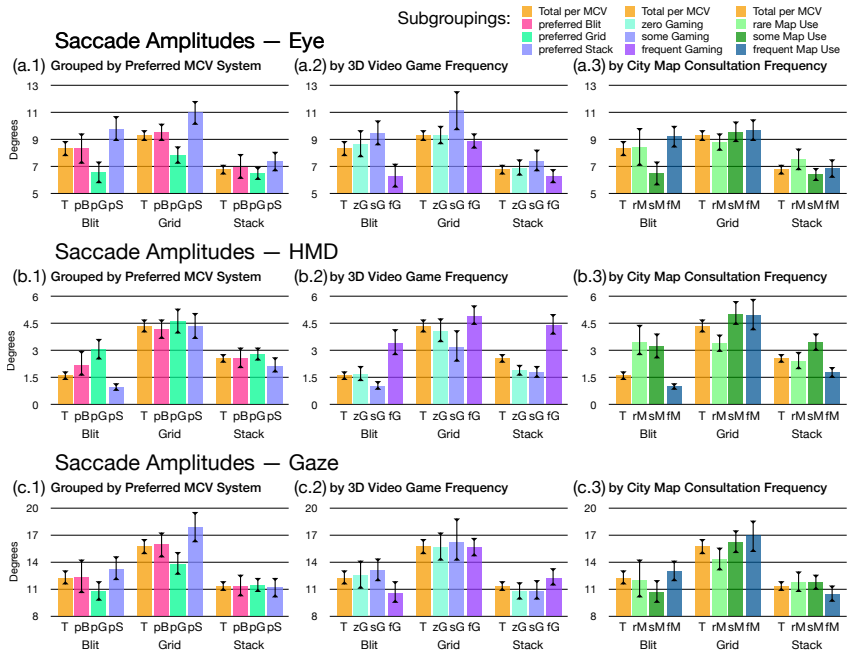


Fig. 6.14: Average saccade amplitude measurements for each MCV system (blit/grid/stack): split by the eye (or viewport) component (**a**), the head movement (or HMD rotation) component (**b**), and taken together to determine the overall gaze (**c**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3).

More substantial differences between the MCV systems and the participant subgroups than those visible from the interaction data (Figure 6.12 and Figure 6.13) could be obtained by investigating their average saccade amplitudes (Figure 6.14), which consist of the eye movement component (a) and the head movement, as tracked by the HMD's position (b), to obtain insight into the overall movement of a participant's gaze (c) when interacting with the systems. This visuo-motor behavior is further illuminated by an analysis of the average durations of the fixations between those saccadic movements, and the average peak velocities reached during these saccades, as shown in Figure 6.16.

Both head and eye movements, as expressed by saccade amplitudes (Figure 6.14 a, b) were clearly the highest for the grid system, and the lowest for the stack system, which corresponds to the distances between the layers in both systems. The blitting system is in the middle, as there is only one map view, but it is quite large itself. All of the subgroups show significant differences within the systems for eye movement, except with the stack MCV (Figure 6.14 a1, b1): the closeness of the layers appears to have led most participants to similar behavior, whereas the other systems led to differentiation in accordance to the subgroups.

Participants who prefer the stack have had far larger amplitudes in their eye movements in the other two systems: they seem to have used their strategy that works for them in the stack, which could be a frequent comparison of far-apart regions or different layers to one another, in the other systems as well, where it would require much larger eye movements. This additional distance may be a factor in their choice of ranking the "larger" systems as worse in terms of legibility.

Looking at the saccade amplitudes as subgrouped by video game playing frequency (Figure 6.14 a2, b2, c2) shows a differentiation between this grouping and that of the preferences in (Figure 6.14 a1, b1, c1). The most frequent gamers move their heads for more and in larger motions than those who play less ((Figure 6.14 b2), and inversely move their eyes far less than those play less (Figure 6.14 a2). These differences cancel each other mostly out, and the overall gaze by the most frequent gamers is much more similar to that of the other participants (Figure 6.14 c2), save for a tendency to have smaller amplitudes when using the blitting system, and larger ones in the stack than the average.

The saccadic movement as subgrouped by preference (Figure 6.14 a1, b1, c1) is however almost perfectly mirrored in the subgrouping by map reading frequency (Figure 6.14 a3, b3, c3), which is interesting, considering how these groups do not overlap, and how much they differ in their interaction patterns (Figure 6.13). Here, participants who consult city maps most frequently trend toward the extremes in terms of their gaze amplitudes: they have the lowest ones with the stack map, and the highest when using the grid, when compared to less frequent map readers. The appear to be more strongly adapting their behavior to the system they are using, being probably already more familiar with how to read maps.

Saccades and their components are also visualized in Figure 10.9, Figure 10.8, and Figure 10.10 in the annex (section 10.3) as polar plots of the density probability distributions. The figures are organized using the same groupings as the bar charts discussed in this section. In addition to confirming some of the above findings visually, they also include information on the direction the saccadic movements take. An excerpt is shown in Figure 6.15, just for the video game frequency subgroups, in which the distinct shapes of the three systems become apparent.

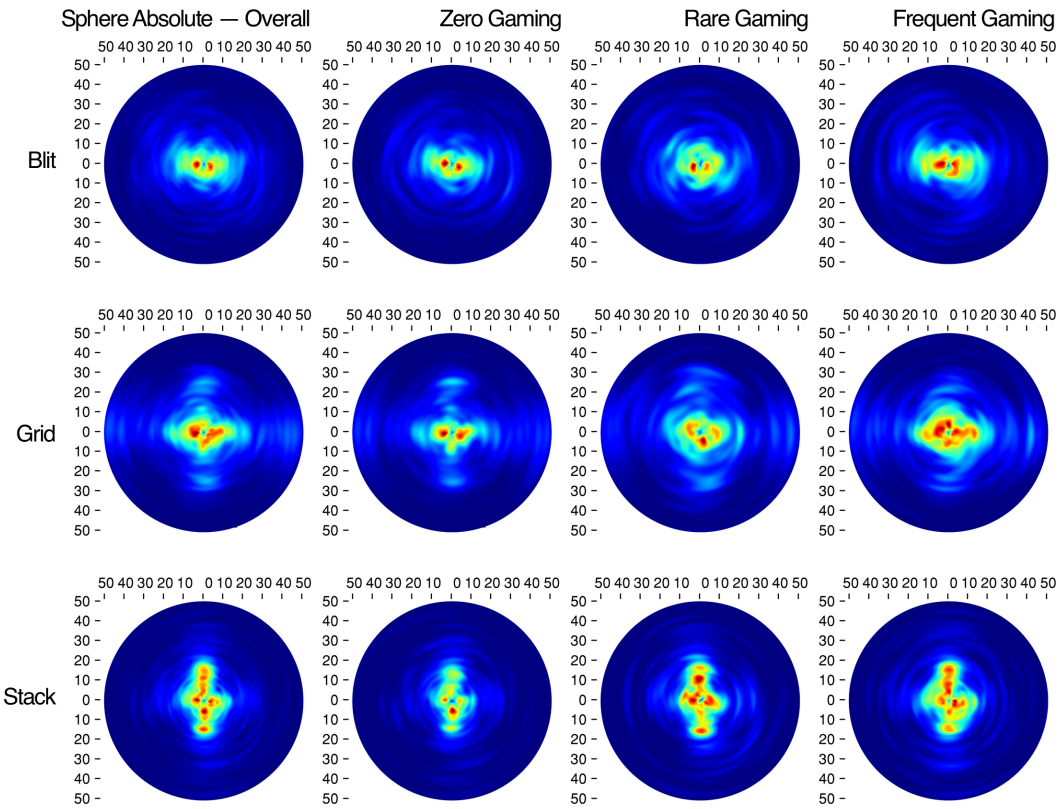


Fig. 6.15: An example of a polar plot, showing the density probability distributions of absolute saccades—the gaze, as projected on a sphere—in degrees, for each MCV system (blit/grid/stack) and subgrouped by the participant's frequency of video game play.

These polar plots inform about saccade behavior for the eye, the head, and the overall gaze movements (the combination of head and eye movement). The distance to the center encodes amplitudes (in degree of field of view), while the angular dimension communicates about two different aspects of saccades. One is absolute angle: the direction of the saccade relative to a horizontal plane (in this 3D case: the longitudinal axis). This information tells the reader how a participant moved their gaze relative to the environment: did they make more saccades vertically or horizontally? On the other hand, the second set of data corresponds to relative direction of saccades: the angle between two saccades [Dav+19].

Relative angles inform about short-term temporal dynamics of saccades: did participants plan a lot saccades in the same directions, and are there signs of back-and-forth motion between several regions of interest? *Forward* saccades (located in the east quadrant) show saccades going farther towards the same direction, possibly to reach a new region of interest—it often represents a shift in attentional target. *Backward* saccades show a behavior where a new saccade is made toward the starting position of the preceding saccade (located in the west quadrant), possibly going back to it for closer inspection or further analysis.

Through the plots of the absolute angles it can be seen that participants produced mostly horizontal saccades when exploring the maps in the blitting system. Less head motions show in this condition, because there is no need for longer saccades to fixate another layer—it can be switched in place. The grid system resulted in a horizontal bias as well as a smaller, but interesting bias showing saccades directed in all cardinal directions, reaching amplitudes of approximately 20 degrees. The same horizontal bias can be observed as in the blitting system; the cardinal bias probably indicated shifting of attention between layers with large horizontal and vertical saccades. Finally, the stack system elicited a substantial amount of vertical movements by eye and head alike—this behavior is necessary to shift from one layer to another located below or above.

When considering relative saccade directions, eye and overall gaze data shows a clear bias towards backward saccades for the blitting system. It is indicative of movements between regions of interest to analyze the content of a layer in help of making a decision according to the task. The forward bias is minor here, as there is little need to make several saccades in the same direction, as one would to orient one's away from the map; additionally, the maps are not so large in the field of view that they require large shifts of attention. In the grid and stack systems both, strong backward and forward biases can be observed. These backward biases can be demonstrative of fine analyses of a map region, while the forward ones account for map exploration. In all three MCV systems, the head rotates mostly further in the same directions (forward relative motions)—it probably does not participate significantly in the fine analysis of a layer's contents.

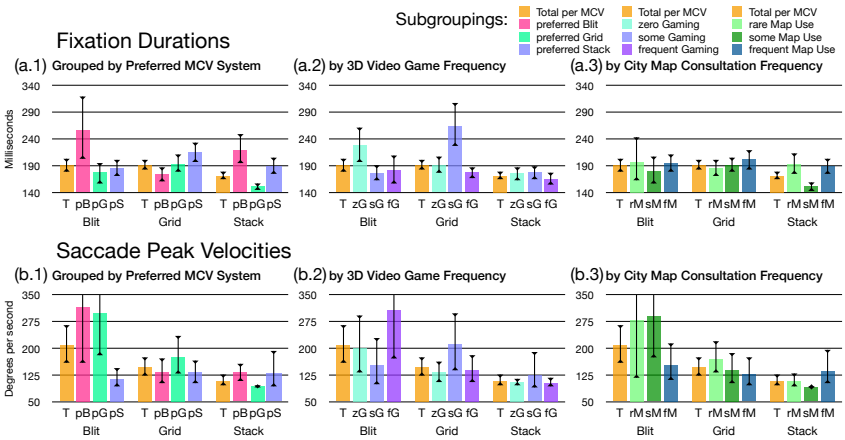


Fig. 6.16: Further visuo-motor behavior measurements for each MCV system (blit/grid/stack): the average fixation durations between saccades (**a**), and the average peak velocities reached during saccades (**b**) (upper confidence intervals cut to increase legibility of smaller values). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3).

Besides the saccades discussed above, the time *between* saccades can be analyzed, in terms of the durations of such *fixations*. When comparing fixation durations across the three MCV systems (Figure 6.16 a), the stack system has significantly shorter ones than

the comparable blitting and grid systems—the closeness of the layers may encourage viewers to linger less on any point and switch their gaze around more, as also seen in the higher number of gaze changes between map layers in that system (Figure 6.12 b).

The differences among the subgroup by system preference show a tendency of participants who prefer the stack to linger longer on points than those who prefer the grid, and those preferring the blitting system having a higher variance overall and lingering much longer on their preferred system. Frequent gamers (Figure 6.16 a2) show the least variation from system to system, with generally low fixation durations—they seem to adapt their search behaviors to each system.

(Figure 6.16 b) also shows the peak saccade velocities. These are similarly the lowest in the stack MCV, and the blitting system having the highest ones, along with the largest variability, as measured by both the confidence intervals and the differentiation the different subgroupings show there. Participants who prefer the stack are also very consistent in their peak velocities across the systems (Figure 6.16 b1), as are those who consult city maps with the highest frequency (Figure 6.16 b3), whereas members of the other subgroups display a much higher variance in this regard; no subgroups of gamers show this kind of stability ((Figure 6.16 b2) at all.

These two groups (the most frequent map readers, and those who prefer the stack) do not exhibit this consistency with other measurements, where their values often vary greatly, such as saccade amplitudes, or the number of their gaze changes between layers. This could point to a higher confidence during map reading in these two groups, regardless of which system they use.

6.3.4 Feedback and Comfort

Besides these measurements and the pairwise comparison questionnaires, participants in the study were also asked to provide free-form feedback at the end of a session regarding each system, in which each MCV system elicited positive and negative remarks. The blitting system received complaints regarding the amount of interaction required to compare layers, which made those comparisons take more time and be more difficult — this is well corroborated by the above findings. Its bigger map view however was welcomed as making analysis of any single layer easier than with the other systems.

The simple arrangement of the grid system was mostly positively remarked on for providing a good overview over the layers. However, the size of the views caused problems in two ways: there were complaints about the smaller size of these vies compared the blitting system, as well as about the large head movements that were required to compare them. This again tracks with observations made above from measured data.

In contrast to these two systems, the stack MCV mostly received suggestions for improving particular parts of it, such as the specific arrangement of the layers regarding their size, tilt, and position; with different and sometimes conflicting suggestions from different users. Multiple participants pointed out their wish to be able to rearrange the layers in a different order. More general feedback included the stack offering a "practical ensemble view" of all layers, and it being the system that allowed the fastest comparisons between them.

Interestingly, the preferences of the participants as expressed through their pairwise comparisons in the questionnaire were not as correlated to their free-form feedback as could have been expected. Some used that text field to name their favorite system, and it did not always correspond to its rank as determined through the PCMs. A connection could be found however between the 3D video game playing habits of participants and the amount detail of their feedback: those who played the most were generally more likely to offer constructive criticisms, often detailed. The frequent gamers also exhibited an inclination for positive comments on the stack system, or at least on ways for how to improve it, having shown promise to them. They thus appeared to be most interested in it, perhaps because of its novelty.

As is customary for studies involving VR, the classic *Simulator Sickness Questionnaire* [Ken+93] was used to look for indications of cybersickness to evaluate the effects of the virtual environment and the systems within them on the well-being of the users. As a tool developed to assess sickness symptoms in flight simulators, it has also been valued for that purpose in VR research. However, it has recently (after the time of this study) been shown to be lacking in its psychometric qualities for measuring cybersickness in commercial VR applications, and updated methods have been proposed [SB20][BWK20].

The SSQ was administered at the beginning of each session, and regularly after task completions. Most participants reported no significant symptoms. Most frequently, "none" was given for all symptoms, and only occasionally "slight." Only one participant complained of "severe" symptoms during the test, and these lessened as the session went on. Cybersickness typically manifests itself from the vergence-accommodation conflict, or a mismatch between visual and vestibular stimuli—when the camera, i. e. the user is shown to be moving while actually remaining seated or standing. This is not the case in any of the MCV systems, as the only movements of the camera or the surrounding room are direct results of the participants' own movements, registered through tracking of the headset.

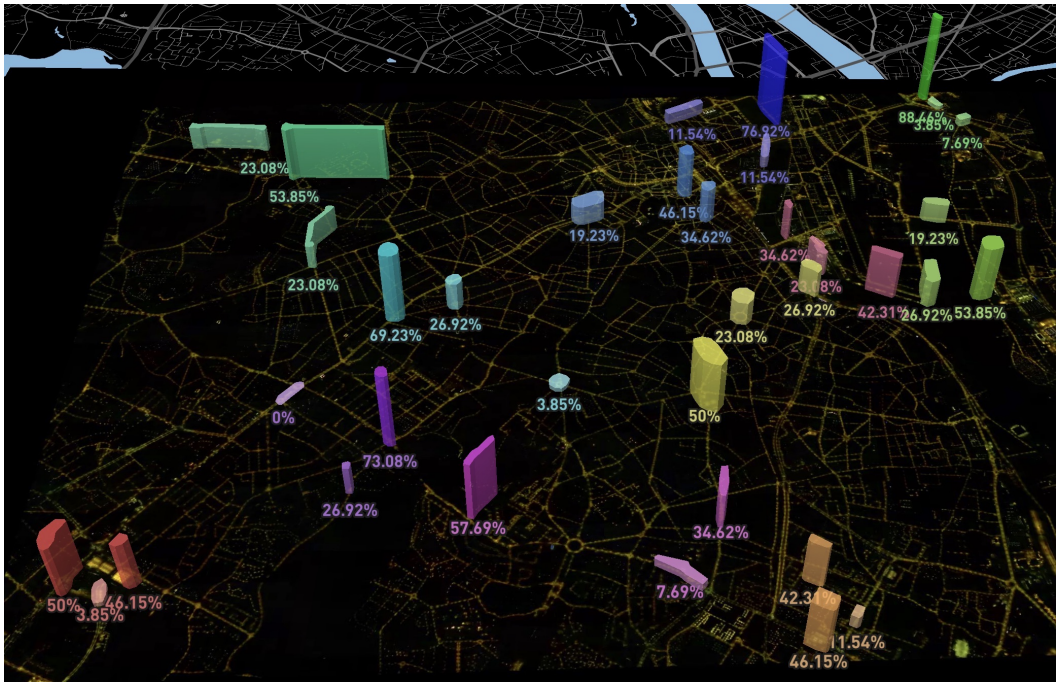


Fig. 6.17: An overview over the 12 scenarios (color-coded) and the distribution of choices participants made between the three candidates in each.

6.3.5 Scenario Choices and Explanations

The other part of the study, after the analysis of the interaction with and perception of each MCV above, is the urbanist use case itself. This relies mostly on the other measurements taken during each scenario: the choice of candidate area, and the free-form explanation of that choice, as well as which layers participants were actually looking at to make their choices. Figure 6.17 shows a distribution of the choices for each region mapped directly on a map of the city, which can give a quick overview. Deeper insights require however an analysis of the written answers; classifying each one by which map layers were referenced and how often, and if the choice and answer made sense in the context of their scenario from the standpoint of an expert urbanist.

Tallying these measurements (relative time spent gazing at a layer and how much it was mentioned in explanations) up over each participant's answers for each scenario (Figure 6.18) allowed for insights into which regions elicited which responses, and how they were argued. The same data could also be used to analyze how individual participants reacted (Figure 6.21).

Analysis of variance shows that these measures vary extremely from person to person ($p < 0.01$ for all data layers), as is expected, and slightly but still significantly from scenario to scenario, as Table 6.1 shows. The relative time spent gazing at different layers was

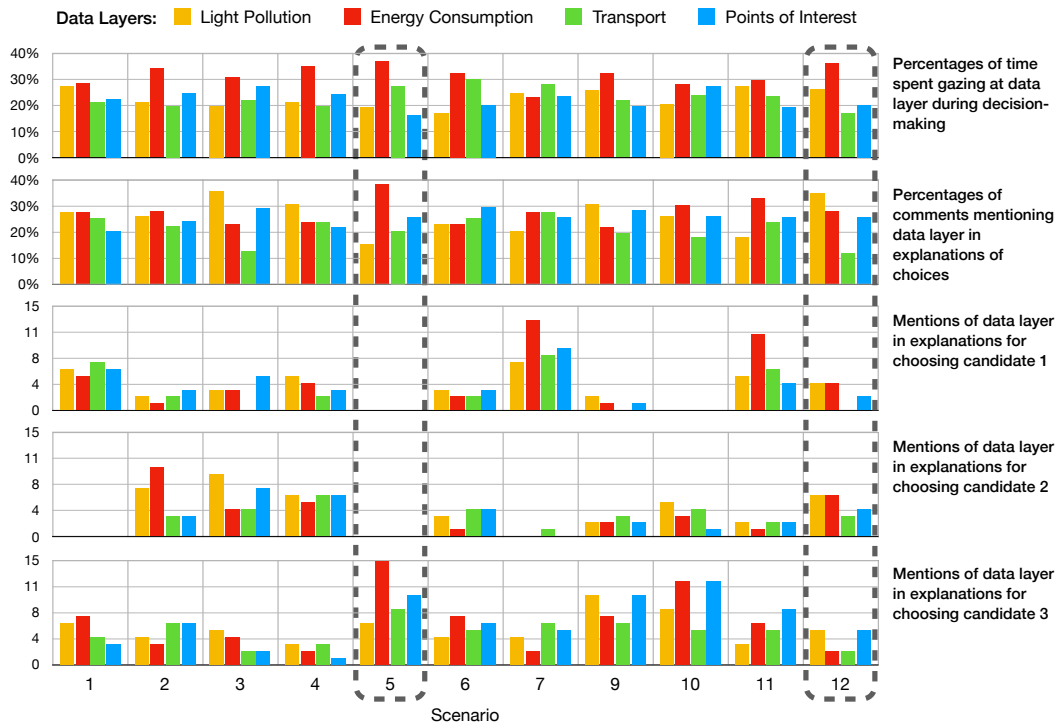


Fig. 6.18: Charts showing data for each scenario: gaze information, candidate choices, and layer mentions in explanations. Scenario 5 and 12 highlighted.

highly affected, especially on the transport (NT) and POI layers. The MCV systems used to display the data however had no significant effect on either of these.

Measure	Variable	LP	EC	NT	POI	Total
Gaze	System	0.15	0.79	0.12	0.22	0.40
	Scenario	0.21	0.31	**0.00**	**0.01**	**0.02**
	Sys:Sce	0.72	0.98	0.10	0.62	0.79
Mentions	System	0.47	0.21	0.92	0.96	0.90
	Scenario	**0.05**	**0.04**	0.09	0.94	0.85
	Sys:Sce	0.96	0.07	0.58	0.62	0.77

Tab. 6.1: Effects of the system used, the scenario presented, and their interactions on time spent gazing at different data layers and in total, and on the number of times one of the data layers was mentioned (Mnt.) in explaining a choice (p-values).

To illustrate the process of the urbanist's analysis of the results for the qualitative study on urban lighting and the level of insight that could be extracted from the data gathered through this comparison of MCVs, contrasting examples of two participants and two scenarios are presented below.

Scenario 5 (Figure 6.19) is chosen as an example because it has the strongest "unbalanced" choices for which candidate area was chosen to improve the lighting situation, with most participants (23 out of 26) choosing candidate 5:3: a primary road axis with crossroads, high energy consumption, high light pollution, low nearby activities, and a presence of public transport. Two participants selected candidate 5:2; a primary road axis with a

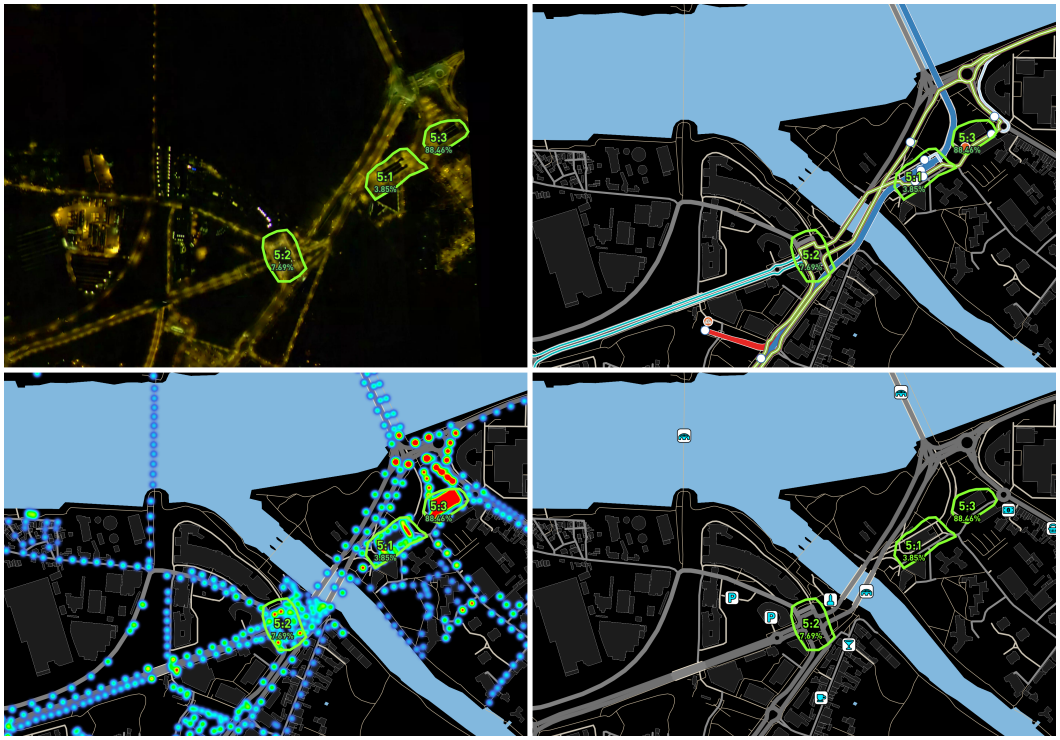


Fig. 6.19: Clockwise from top left: light pollution aerial orthophoto, nocturnal transport network, points of interest, and energy consumption heatmap layers of scenario 5, with candidate area outlines.

roundabout, high energy consumption, low presence of night activities, a presence of public transport, and moderate light pollution. Only one participant picked 5:1, which shows a tram station, high energy consumption, low nearby activities, and moderate light pollution.

The configuration of candidate area 5:3 (at the top-right corner of the layers in Figure 6.19) could be explained by the layout of the city: the area is a covered car park close to one of the most important public transportation stations. It is brightly illuminated, which the prevalence of energy consumption, but a roof over part of that parking area limits a reflection of that light to the sky, which explains the dark square on the light pollution layer.

Layers mentioned in explanations: The 3 participants who chose candidates 5:1 or 5:2 did not mention any of the data they saw when explaining their choices, while the 23 who selected 5:3 mostly referred to the energy consumption (38%), and then in decreasing order activities (26%), transport (21%), and finally the light pollution layer (15%) when arguing their choice, as seen in Figure 6.18.

Layers observed: While selecting candidates in this scenario, the participants spent most of their time (37%) inspecting the energy consumption layer. The other layers are ranked by this metric as follows: transport (26%), light pollution (19%), and lastly the points

of interest layer (16% of time spent gazing at it). The order of how much a layer was looked at differs from how much it influenced the decisions as measured by the number of mentions in the free-form explanations: while the energy consumption layer remains most important in both, the points of interest are least gazed at, while being second-most often mentioned.

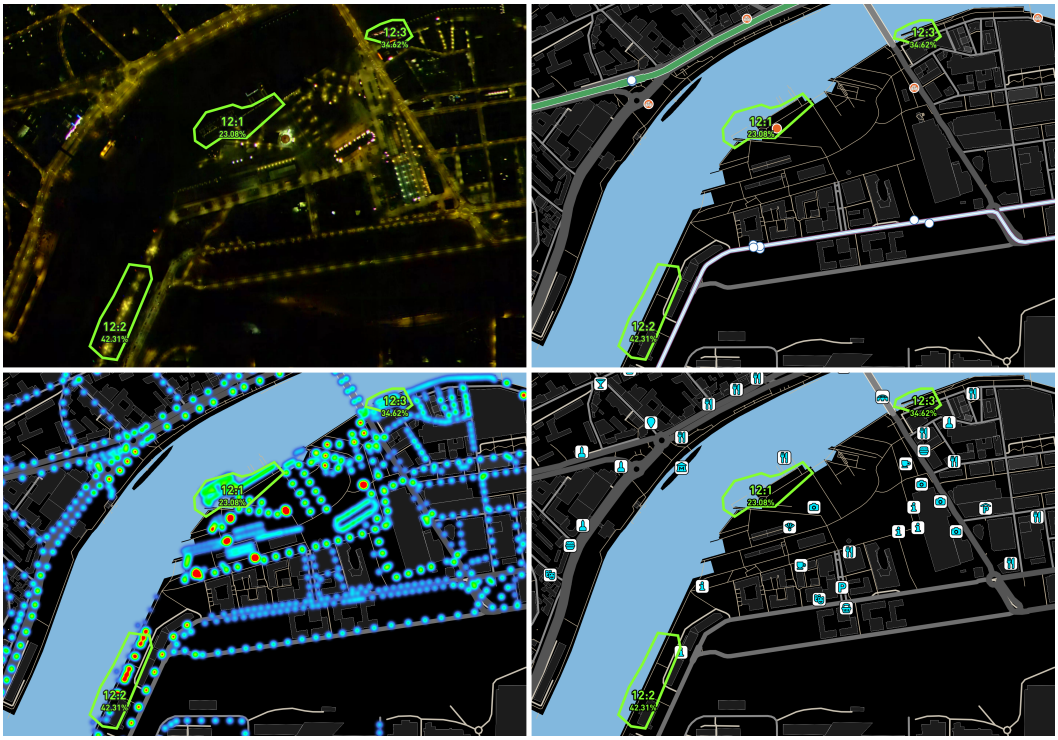


Fig. 6.20: Clockwise from top left: light pollution aerial orthophoto, nocturnal transport network, points of interest, and energy consumption heatmap layers of scenario 12, with candidate area outlines.

Scenario 12 (Figure 6.20) was chosen as an example in contrast to scenario 5 because it hast the most balanced choices among its candidates. 6 of the 26 participants chose candidate 12:1, which is close to a restaurant, with moderate energy consumption, distant from transport, and with low to no light pollution. 12:2 was chosen by 11 participants, is similarly located close to a restaurant, but with localized, very high energy consumption, strong light pollution, and slightly closer proximity to public transportation. The remaining 9 participants selected 12:3, similar in all regards to 12:1, but with more light pollution. All candidate areas are also bordering on a river.

Layers mentioned in explanations: The choices of candidates 12:1 and 12:2 were mostly explained by mentioning the light pollution and energy consumption data (Figure 6.18). When candidate area 12:3 was chosen, the two main explanations were mentioning light pollution and points of interest, and to a much lesser degree the other two layers. Overall, this scenario's choices were motivated mostly by light pollution (35%), closely followed by energy consumption (28%) and points of interest (26%), and to a much lesser degree by the transport network (12% of mentions in the explanations).

Layers observed: As in scenario 5, the most attention was given to the energy consumption layer (36%), at the expense of the transport network (17%), which fell behind light pollution (25%), and received less attention than the points of interest (20% of time spent gazing at its data layer).

Links between scenario 5 and 12: In scenario 5, the vast majority of participants chose candidate 5:3. Highly prominent on the energy consumption layer, it seems to have caught the attention of most participants, although it is directly linked to transport and is not responsible for much light pollution. In scenario 12, the choice is not so clear: each candidate is on the bank of a river, and some participants in their explanations referred to the introduction text, where fish are mentioned as victims of light pollution. Energy consumption is the most observed data layer, even if it only takes second place in their mentions. Instead, light pollution is the most commented data layer, even if candidates 1 and 2 show less of it in comparison to candidates in scenario 5. Perhaps the dark area of the river created a visual contrast that was driving the choices. In both scenarios, the most consuming area "wins:" candidates 5:3 and 12:2 both contain the highest density of energy consumption in their area.

Whether balanced or not, the ratio of observation percentage, comment percentage, and the candidate selection itself appears to highlight an orientation of the participants towards the candidates where energy consumption is highest. This choice could related to an understanding of the task, which was introduced with a focus on reducing energy consumption and light pollution of public spaces.

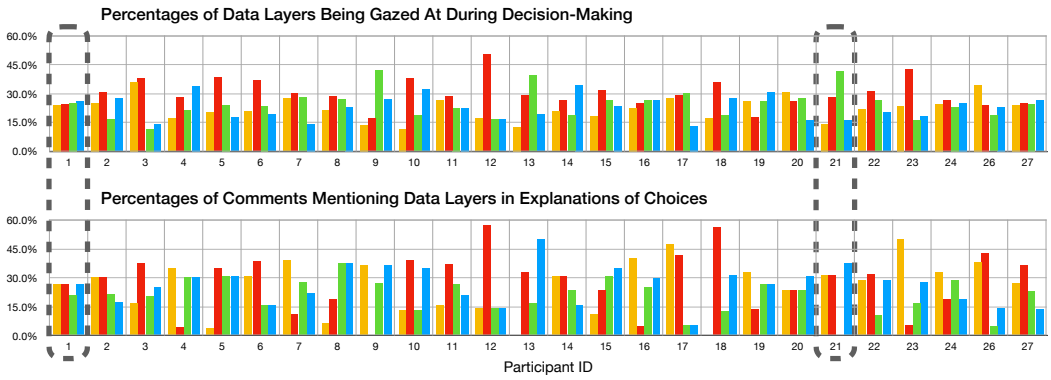


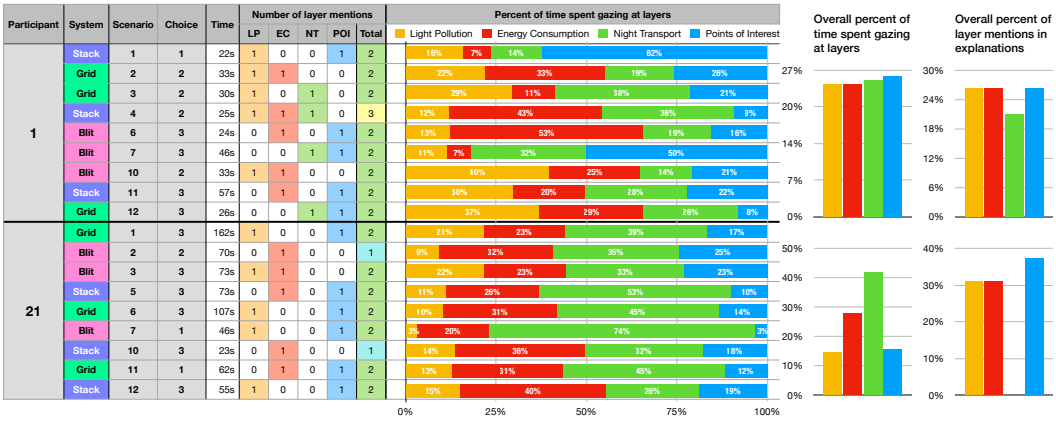
Fig. 6.21: Overview of the gaze and explanations of each participant. Participant 1 and 21 highlighted.

Another purpose of the qualitative analysis of the 26 participants was to observe if their profiles and backgrounds as seen in Figure 6.7 and Figure 6.8 have an impact on their choices in this study. These factors have already been shown to influence and explain preferences for one MCV over another and their performance and behavior in them in the previous sections (section 6.3.1, section 6.3.2, section 6.3.3). Also of interest to the urbanist analysis is to investigate possible links between the attention of participants to the data layers as measured by their gaze and the explanations they gave for choosing

candidate areas, charted in Figure 6.21. An example analysis on participants 1 and 21 follows below.

Participant 1 is an 18-year-old man, who spends about 10 hours a week outside after sunset (Figure 6.7 and Figure 6.8 in green.) He is a student and has average familiarity with maps and the city shown in the maps. His gaze and explanation percentages are well-balanced, in that overall, he spent roughly equal amounts of time observing each data layer, and referenced each of them equally often when explaining his choices (Figure 6.22 on top).

He regularly mentions the criterion of energy consumption in his explanations, which he describes almost every time as "huge." He is interested in the attractiveness and accessibility of the candidates he selects (1:1; 2:2), because, according to him, lighting is synonymous with life, and a place without lighting will therefore be rendered "dead" (6:3). He consolidates his position by choosing for example a candidate he describes as "very attractive and not well lit, which in itself is contradictory" (3:2). Finally and despite the aim of the study, he proposes twice to *add* more lighting to the candidates he selected, in order to improve their attractiveness (2:2; 12:3). For this participant, the notion of public lighting seems decisive in the nightlife that he intends to find in a place, especially if it is easily accessible thanks to a good luminous installation.



Participant	System	Scenario	Choice	Time	LP	EC	NT	POI	Total
1	Stack	1	1	22s	1	0	0	1	2
	Grid	2	2	33s	1	1	0	0	2
	Grid	3	2	30s	1	0	1	0	2
	Stack	4	2	25s	1	1	1	0	3
	Blit	6	3	24s	0	1	0	1	2
	Blit	7	3	46s	0	0	1	1	2
	Blit	10	2	33s	1	1	0	0	2
	Stack	11	3	57s	0	1	0	1	2
	Grid	12	3	26s	0	0	1	1	2
21	Grid	1	3	162s	1	0	0	1	2
	Blit	2	2	70s	0	1	0	0	1
	Blit	3	3	73s	1	1	0	0	2
	Stack	5	3	73s	0	1	0	1	2
	Grid	6	3	107s	1	0	0	1	2
	Blit	7	1	46s	1	0	0	1	2
	Stack	10	3	23s	0	1	0	0	1
	Grid	11	1	62s	0	1	0	1	2
	Stack	12	3	55s	1	0	0	1	2

Fig. 6.22: Analysis of the two selected participants (1 and 21), showing differences in choices, explanations, and visual attention as measured by time spent gazing at different layers.

Participant 21 is a 38-year-old woman who spends around 2 hours a week outside after sunset (Figure 6.7 and Figure 6.8 in blue). She is an employee and has a strong familiarity with maps and the city of Nantes. From her written explanations of the candidate choices a good understanding of each data layer could be observed, such as a strong differentiation between produced light and consumed energy. Her explanations focused mostly on the relation of the public lighting to points of activity, and its energy consumption.

Her behavior during her session can be characterized as unbalanced with respect to her gaze attention and explanations not being aligned to each other (Figure 6.22 bottom). She spent a large amount of time observing the transportation layer, but without mentioning it at any time. It can be assumed that this layer was mostly used for orientation in the city, to place the shown data in known contexts. Although she mentioned points of interest the most in her answers, that layer is along with the light pollution layer the one she spent the least amount of time inspecting.

In the candidates she chose, a focus on nightlife and energy can be observed. She placed a lot of importance on reducing light pollution, while also concentrating her comments on the subject of activities. Unlike participant 1, she focuses on places that are highly lit and yet devoid of activities (1:3; 6:3; 7:1; 11:1; 12:3), which should therefore have their lighting reduced. In the same way as candidate 1 she based this on the importance of energy consumption, which she regularly poses as the main argument (2:2; 3:3; 5:3; 10:3; 11:1).

6.4 Discussion of the Study

The impact of which MCV variation was used had apparently no bearing on the understanding of the data and the answers given, as those depended much more on the data shown and a participant's behavior. It can be seen as a success in making all variations comparably efficient in displaying information, and thus separating the evaluation of their usability against one another from their usefulness for the presented study in urbanism.

The participants in this study as well as the experts consulted during the implementation and pre-test phase have both expressed their enthusiasm for and voiced complaints against each of the three tested systems. By the metrics of this study, the stack MCV could be said to have slightly prevailed in terms of user acceptance, if its visual design is taken into account, but is otherwise mostly comparable to the grid system in terms of map layer legibility and ease of use. Both these spatial multiplexing systems leave the blitting system behind.

However, these subjective preferences do not paint a clear enough picture to allow a declaration of a "winner" among these MCV systems. Neither do the objective eye-tracking metrics, especially since their meaning in this context remains open for interpretation; the speculations in section 6.3.3 remain attempts at just that.

Although there are significant differences in visuo-motor behavior, fixation durations or saccade amplitudes do not directly translate to quality of a system. An exception could

be the amount of head movement, as plotted in Figure 10.9: more movement of a still quite heavy HMD will certainly lead to more strain and potential exhaustion, if continued over longer periods [Yan+19]. In this instance, the stack system clearly wins over the grid system in this regard, most likely due to its more compact layout as hypothesized.

The stack's layout is also likely to have caused the higher number of recorded changes in gaze from one map layer to another in the stack: to switch attention from the bottom layer to the one on top, the intermediary ones need to be crossed, thus inflating this number over that of the grid system. Even so, the number of gaze changes while using a system does not directly speak to the quality of it; higher numbers could indicate both a deeper process of comparison, as well as a higher state of confusion or disorientation, while lower ones could be indicative of more focus, or of a higher inconvenience of switching to different maps, as is definitely true for the blitting system, which requires an additional controller interaction to do so.

Task completion times were highly variable and show no significant differentiation across the MCV systems, which can be attributed to the differing difficulties and complexities of the scenarios, and to the fact that the participants were not asked to complete the tasks as quickly as possible, but rather provide their best responses, even if that would take more time. The results from section 6.3.5 show this actually happening, as the choices and explanations given by the participants were sensible and useful for urbanist analysis, regardless of which MCV system was used for evaluation of the public lighting scenarios.

This last point can be seen as an indication for achieving one of the design goals set out: implementing all three MCV systems to be as functionally equal and usable as possible. The inherent differences between the systems had no measurable impact on the quality of the insights for the urbanist study attained through them, only on the comfort and performance of the participants while using them. This point also addresses the contribution to conducting studies with multi-layer map comparisons in VR: the setup as reported here has been shown to be suitable for providing valuable insights for research in urbanism. With all MCV systems performing equally in this regard, the contribution of evaluating the novel stack system thus needs to be more nuanced than to declare it "better" or "worse" than the other two systems it was compared against.

The hypothesis concerning the benefits of the stack MCV's layout—the closer proximity of the map layers while maintaining legibility—does bear out with the above analysis for the recorded measurements, but it appears not to be the only factor in determining its performance or acceptance among participants. Indeed, the participants themselves have been found to be the deciding factor in these matters: their characteristics, as seen through questionnaires about their habits and as measured by their visuo-motor behavior are significantly influential in which system they prefer or how they interact with it.

To provide an optimal experience for similar map comparison tasks, the results indicate that just optimizing and improving any one MCV system would not work universally. Different groups of the population appear to prefer opposing aspects of the different systems, depending on their experiences, expectations, and behavior. A clear example for this were the participants who frequently play 3D video games, and thus have probably greater experience with interaction and navigation in virtual environments. They showed a higher inclination to interact more deeply with the blitting and the stack system, both more "advanced" than the grid, in either requiring the use of controls to switch layers, or in being arranged in a more unusual way than the simple grid. This grid, conversely, was by far the most popular system with participants whose video game frequency was the lowest, or nonexistent.

6.4.1 Limitations

This study was designed to evaluate each MCV concept on its own within different scenarios, without combining them, so as to better isolate their effects on user behavior and performance. Furthermore, a comparison between MCVs and a single map view containing all data by itself was omitted—such a comparison would fall more into the domain of cartography; of how much information could be put into one map. The use case was specifically designed so that a single map would not be practical, and some kind of splitting of the information layers had to occur, so the focus here is instead on comparing methods of such splitting. The methods were also supposed to be a simple as possible, as opposed to selective schemes and other, more advanced methods of bringing more information into one single view.

In this study design, the number of data layers was also kept at a minimum so as to maintain an "even playing field" for all systems. This number (four) provides an ideal case for both the grid and the blitting systems: all layers could be mapped to the cardinal directions for switching between them, and put in the corners of a rectangular frame for a clean grid view. Higher numbers of layers would complicate such arrangements in both these systems, and potentially necessitate smaller views in the grid, while having no such obvious downsides for the stack system other than requiring a slightly stronger tilt.

Following this reasoning, stacking could accommodate a larger number of layers in the same visual space, or present the same number of layers closer to each other. Even a moderate inclination of the stack layers results in large savings of vertical field of view, thus having a clearer advantage over the other systems, e. g. by having more layers in the same field of view, or increasing the ease of map comparisons. Tilting the layers can increase the amount of horizontal space that they take up, depending on where that relative inclination occurs, by bringing the lower edges closer to the viewer. Figure 10.11 and

Figure 10.12 in the annex illustrate the space savings of stacks with different inclinations over coplanar arrangements and show possible hybrid systems.

The limiting factor of how much vertical space can be saved with the stack system is the decreasing legibility of the views' content with increasing tilt. In addition to Figure 6.3 for four layers, Figure 10.11 and Figure 10.12 in the annex show different arrangements of the MCV stack in comparison to grids. While colored dots still appear legible at a very small viewing angle, text labels and icons begin to lose legibility with higher distortion. Display resolution is also an obvious factor in this, but in any case further research is warranted into directly comparing the legibility of different kinds of information across inclinations and distortions, and into exploiting the virtual nature of these map views to e g. elevate and billboard labels and icons for greater visual clarity.

Similar to limiting the number of map layers for the study, the interactions with the MCVs were also limited, especially with the technically very customizable stack system. The parameters dictating the shapes and sizes of the layers, their positions and viewing angles were all kept static, in order to ensure consistent and thus more comparable conditions for all participants. The initial choice of where each map layer is situated within one of these MCV systems may thus lead to a less than optimal positioning for a particular user or scenario. Free-form feedback from the study showed multiple wishes for exactly that kind of flexibility.

Another limitation imposed in the name of greater consistency was to have the participants sit at a table for the study. This limits the utilization of the virtual space to its fullest extent, e. g. practically none of the physical navigation and the associated distributed cognition that other proposals, such as that by Mahmood et al. [Mah+18] (seen in Figure 4.6) espouse, and that especially spatial multiplexing MCV systems could stand to benefit from greatly.

6.5 Conclusion

Basing design choices and arguments on previous work regarding composite visualizations and map comparison techniques, the stack system was introduced—a novel MCV system that aims to take advantage of the virtual and interactive environment current immersive technologies offer, such as the large work spaces, stereoscopic presentation, and spatial tracking. A task was designed to evaluate it in a user study against comparable, more traditional MCV methods. This study contributed to the investigation into extending the MCV concept to immersive environments (specifically VR in this case), with a focus on geospatial visualization, and using them for comparisons of more than just two layers of data. It further directly integrated an actual use case in urban design

research in a hybrid study design. At time of this writing, a combination oven even just a subset of these inquiries has not appeared to have been investigated rigorously, making this study possibly the first of its kind.

Using eye-tracking technology, each individual participant's behavior could be measured and analyzed in how they read the individual map layers and reacted to the different MCVs on a subconscious level. Correlating these very fine-grained measurements with the expressed preferences of the participants for certain systems, as well as with their habits and experiences, conclusions could be drawn as to which kinds of MCVs are better suited to different types of users and why. Furthermore, this kind of analysis encompassing many different measurements has indicated that no singular metric is sufficient to allow for a meaningful comparison between such system.

Consequently, while the hypothesis that the novel stack system has certain advantages over comparable MCV methods for map comparisons could be confirmed, the evaluation did not result in a universal picture. For instance, participants most adept at interacting with virtual 3D environments and with the fastest eye movements were found to prefer the stack over other MCV systems, which could point toward it being a "better" system if fully taken advantage of, but those characteristics can not be expected of all users. Instead, a choice of different MCV methods, and perhaps the freedom to alter their arrangements should be offered to users for actual tasks, as the effects of MCV system on behavior were significant, and user preferences differed strongly. All three methods were however equally useful in productively completing the task from the chosen use case in citizen participation in questions regarding public lighting, where the produced insights were valuable for the urbanist research into that topic.

In addition to hitting the design goals for an MCV that takes advantage of VR, all three goals of the study as set out in section 6.2.1 appear to have been met: differences between the perceptions of three implemented MCVs have been found and could even be attributed to differences between the participants, as determined through questionnaires and analysis of eye tracking data. Immersive MCV systems were found to be useful in an actual participatory study in urbanism, and the specific layout did not impact results there. A large portion of the participants showed eagerness to take part in more studies of this kind, expressing clear enthusiasm for this kind of use of immersive technologies. Though not directly asked as such, this feedback can be seen as a measure of how *engaging* the immersive MCVs appeared to users, which constitutes a clear "win."

7 Virtual Data Sphere for Inverse Stereographic Projection

Similarly to how idea of the inverse stereographic projection sprang originally from a modification of the MCV stack system, its implementation is built on the same foundation as discussed in the previous chapter (section 6.1.1): Unity 3D as the engine, and Mapbox as an interface for geospatial data. The same hardware for VR was used as well: the first-generation HTC Vive setup. This chapter describes how the spatial deformation described in section 5.3 was achieved in Unity and how it was made interactive, and discusses the resulting distortion and how it compares to existing focus+context methods. Finally, some potential applications of this projection are presented.

7.1 Implementation: Virtual Data Sphere and Other Surfaces

7.1.1 Projection by Shader

For the stack MCV system, and in fact for all other MCVs implemented for the comparative study, shaders were used to frame the available map tiles from the Mapbox SDK into well-defined views (section 6.1.2). They acted as *cutout shaders*: being given the extents of a box (the frame) as parameters, the texture of the parts of the map falling outside this box was rendered as invisible.

Besides determining the look of textures, shaders are also capable of displacing vertices before rendering them according to any set of rules. For polygon meshes that are sufficiently finely tesselated[1], shaders could thus be used to create any of the deformations discussed in section 5. Of course, the same could be accomplished with methods executed on objects in Unity, however that would not benefit from execution on the GPU and the associated massive parallelization and thus performance increase.

[1] Polygon count can be a problem with meshes of maps. Without a digital elevation model, whole map tiles can be output as simple quads, and straight roads as chains of very narrow trapezoids. The Mapbox SDK was edited here to subdivide roads into chunks of a given maximum length, so that they could be bent more smoothly by vertex displacement.

Since Unity version 2018.1, shaders can be created visually, using the *Shader Graph*, but creating and connecting nodes in a graph network. Most shader functions are available as nodes, and custom nodes can be created to execute arbitrary C# and HLSL code. Especially for prototyping and iterating, the visual nature of a graphic layout of functions can be very efficient, as it allows an intuitive overview over the complete flow, from the inputs and properties to the final "master" output.

In the introduction to the inverse stereographic projection (section 5.3.4), it was explained as rays connecting the north pole of a sphere to points on a plane/terrain intersecting it, and using these rays to project the image of the terrain on the sphere (Figure 5.12). This can be directly formulated into an expression suited for a shader, requiring only the position and radius of the sphere, in addition to the already given positions of each vertex.

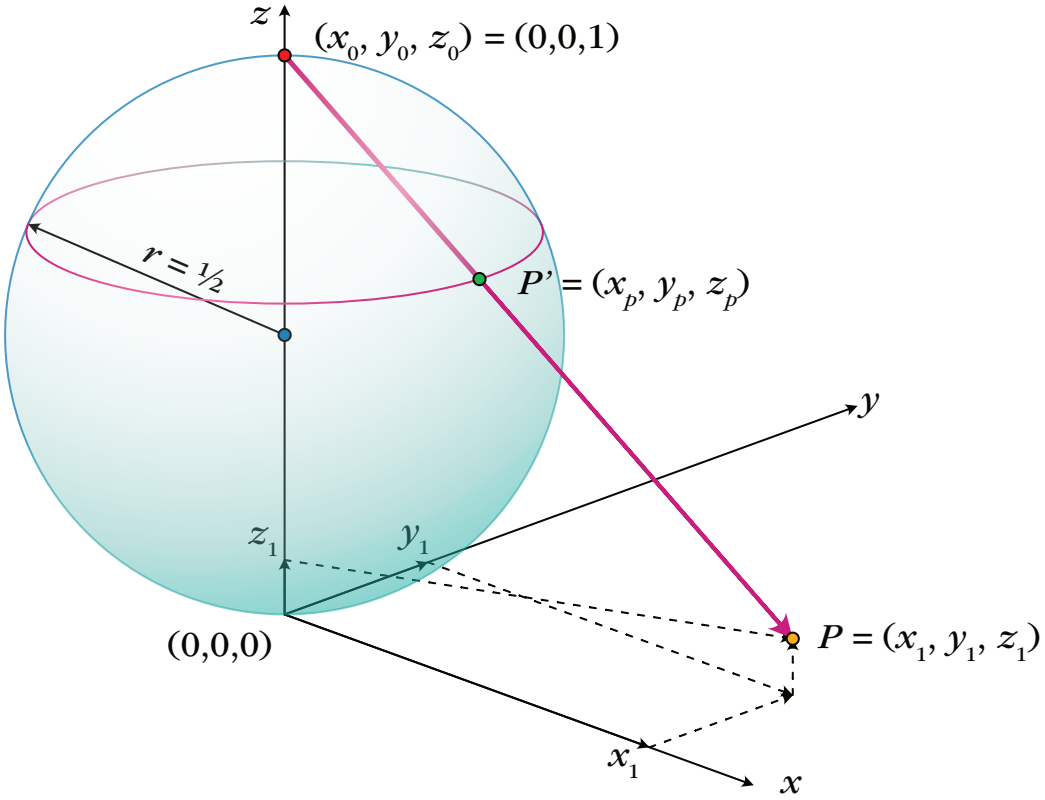


Fig. 7.1: Inverse stereographic projection of a point P to P' on a sphere with a radius of $\frac{1}{2}$, and its center at $(0, 0, \frac{1}{2})$.

The expression can be further simplified if the scene is first normalized: instead of calculating the projected position directly from the two original positions of sphere and vertex, the scene can be transformed to be centered around the south pole of the sphere, and scaled so that the sphere has a diameter of 1. The projected point $P' = (x_p, y_p, z_p)$ can now be determined as being at the intersection between a line going through the north pole, or the projection origin at $(x_0, y_0, z_0) = (0, 0, 1)$ and the original (but normalized) point $P = (x_1, y_1, z_1)$, and the sphere's surface, given by Equation 7.1

and shown in Figure 7.1. The coordinates of this line can be represented by Equation 7.2; generically, and with the projection origin and end point filled in, with the parameter l for all points in between. Equating the points (x_l, y_l, z_l) on the line and (x_s, y_s, z_s) on the sphere results in a single solution for l (Equation 7.3), which when put back into the line's equations (Equation 7.2) gives the coordinates of the projected point $P' = (x_p, y_p, z_p)$ in Equation 7.4.

$$x_s^2 + y_s^2 + \left(z_s - \frac{1}{2}\right)^2 = \frac{1}{4} \tag{7.1}$$

$$\begin{aligned} x_l &= (x_1 - x_0) \cdot l + x_0 = x_1 \cdot l \\ y_l &= (y_1 - y_0) \cdot l + y_0 = y_1 \cdot l \\ z_l &= (z_1 - z_0) \cdot l + z_0 = (z_1 - 1) \cdot l + 1 \end{aligned} \tag{7.2}$$

$$l = \frac{1 - z_1}{x_1^2 + y_1^2 + (z_1 - 1)^2} \tag{7.3}$$

$$\begin{aligned} x_p &= \frac{x_1 \cdot (1 - z_1)}{x_1^2 + y_1^2 + (z_1 - 1)^2} \\ y_p &= \frac{y_1 \cdot (1 - z_1)}{x_1^2 + y_1^2 + (z_1 - 1)^2} \\ z_p &= \frac{x_1^2 + y_1^2}{x_1^2 + y_1^2 + (z_1 - 1)^2} \end{aligned} \tag{7.4}$$

A function that computes the final coordinates according to Equation 7.4 was straightforwardly created as a custom node in the shader graph, taking in the normalized coordinates and outputting their projection. After the inverse of the aforementioned normalization process, the basic inverse stereographic projection is now complete. As an external parameter, this shader only requires the position and diameter of the sphere, aside from textures, colors, etc. Figure 7.2 shows a plane textured with a repeating pattern of circles and perpendicular lines, akin to a "quasi-indicatrix" [Bil15] modeled after Tissot's Indicatrix [Las89] to indicate map projection distortions, and its inverse stereographic

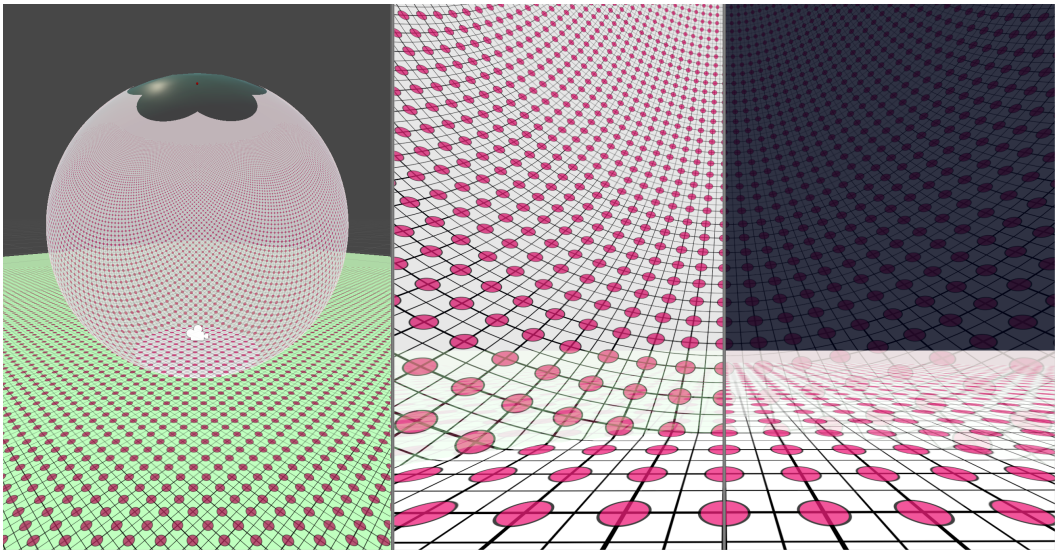


Fig. 7.2: External and internal views of a simple inverse stereographic projection of a flat, textured plane, with the original view superimposed on half of the internal view for comparison. The unchanged plane inside the sphere is opaque and white, outside the sphere it is opaque and green, and the projected image is translucent white.

projection on a sphere intersecting that plane. Also shown is the view from inside it, with the camera being located on the north-south axis of the sphere as in Figure 5.12, with the original view superimposed on one half of it for comparison.

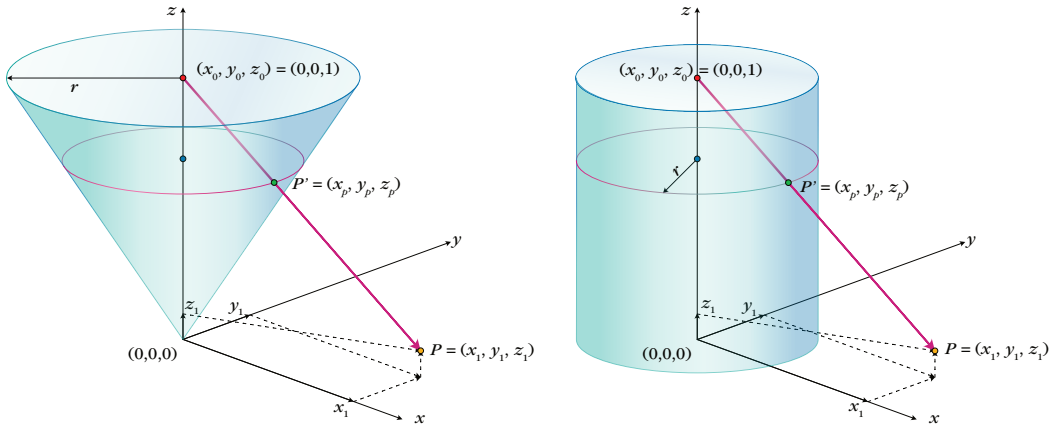


Fig. 7.3: The principle of the inverse stereographic projection applied to the cone and cylinder.

The exact same approach was taken for computing "stereographic-like" projections onto a cone or cylinder, as first discussed in section 5.3.6, and shown in Figure 7.3. By constraining these shapes to their "equivalent" sphere as shown in Figure 5.15, i. e. mandating for them to be tangential to the sphere at the intersection with the ground plane and their projection origin to be the same, their geometric construction also becomes easier. Only an additional normalization step is necessary before (and after) projection on the cone: instead of intersecting the ground, its apex is lifted to the coordinate system origin. The radius and position of the equivalent sphere thus fully dictates the radius of the cylinder, and also the base radius of the cone at the height of the projection.

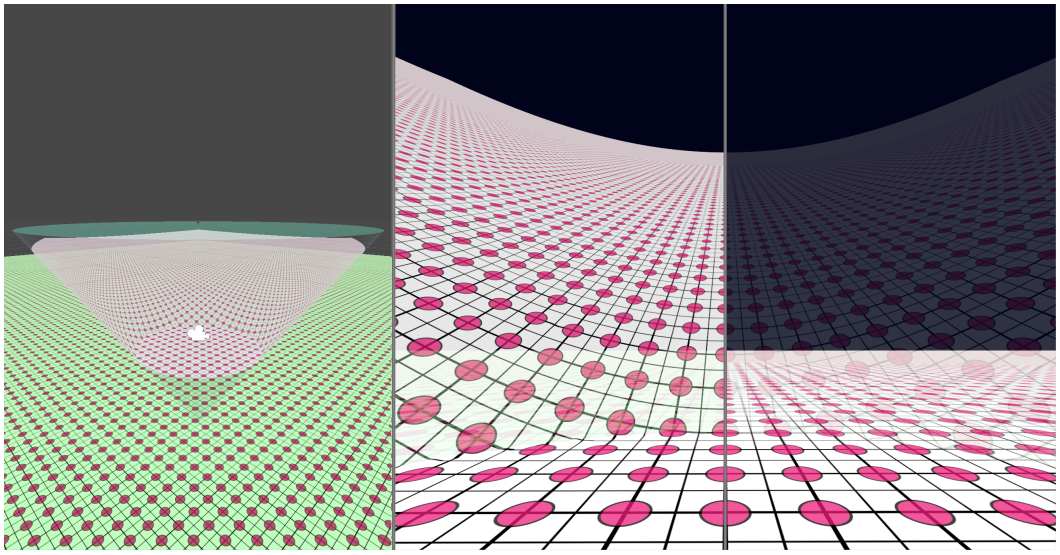


Fig. 7.4: The projection on a cone's surface, as in Figure 7.2.

The cone's surface is given by Equation 7.5. Equating it with the same line (Equation 7.2) from it's projection origin $(0,0,1)$ at its base to the original point $P = (x_1, y_1, z_1)$ with this surface results in the line parameter l shown in Equation 7.6, and thus the coordinates of the projected point $P' = (x_p, y_p, z_p)$ shown in Equation 7.7. Figure 7.4 shows the projection being applied to the same scene as in Figure 7.2.

$$x_s^2 + y_s^2 + \left(z_s - \frac{1}{2}\right)^2 = \frac{1}{4} \tag{7.5}$$

$$l = \frac{(1 - z_1) \cdot r^2 - \sqrt{x_1^2 + y_1^2}}{(z_1^2 - 2z_1) \cdot r^2 - x_0^2 - y_1^2} \tag{7.6}$$

$$\begin{aligned}
x_p &= \frac{(1 - z_1) \cdot r^2 - \sqrt{x_1^2 + y_1^2}}{(z_1^2 - 2z_1) \cdot r^2 - x_1^2 - y_1^2} \cdot x_1 \\
y_p &= \frac{(1 - z_1) \cdot r^2 - \sqrt{x_1^2 + y_1^2}}{(z_1^2 - 2z_1) \cdot r^2 - x_1^2 - y_1^2} \cdot y_1 \\
z_p &= \frac{(1 - z_1) \cdot r^2 - \sqrt{x_1^2 + y_1^2}}{(z_1^2 - 2z_1) \cdot r^2 - x_1^2 - y_1^2} \cdot (z_1 - 1) + 1
\end{aligned} \tag{7.7}$$

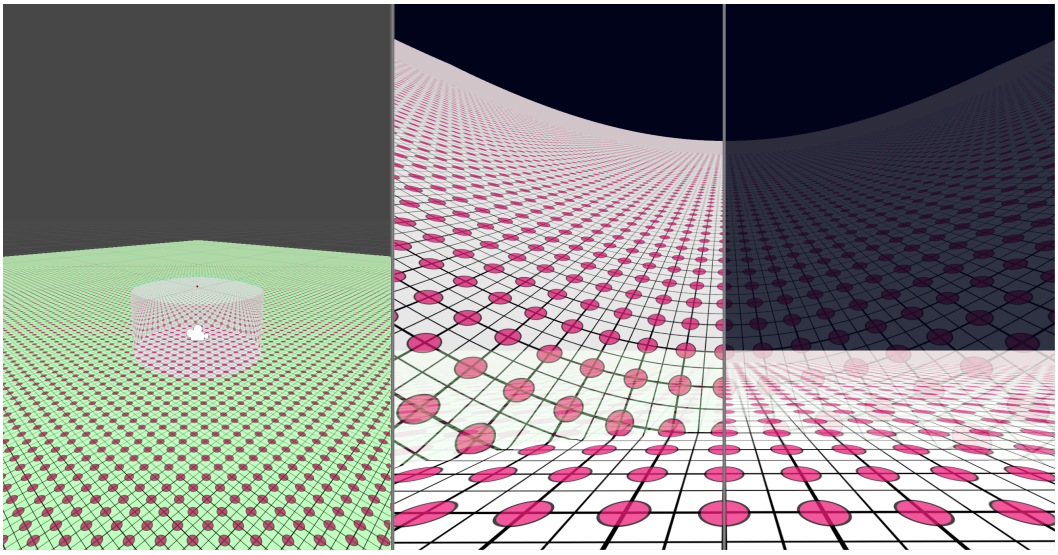


Fig. 7.5: The projection on a cylinder's surface, as in Figure 7.2.

Finally the cylinder's surface is given by the much simpler Equation 7.8. Again equating it with the line (Equation 7.2) from it's projection origin $(0,0,1)$ to the original point $P = (x_1, y_1, z_1)$ with the cylindrical surface results in the line parameter l shown in Equation 7.9. The coordinates of the projected point $P' = (x_p, y_p, z_p)$ are shown in Equation 7.10, and the resulting projection in Figure 7.5, bearing a striking resemblance to that of the cone (Figure 7.4) when viewed from inside, despite their external appearances varying widely.

$$x_s^2 + y_s^2 = r^2 \tag{7.8}$$

$$l = \frac{r}{\sqrt{x_1^2 + y^2}} \tag{7.9}$$

$$\begin{aligned} x_p &= \frac{r \cdot x_1}{\sqrt{x_1^2 + y_1^2}} \\ y_p &= \frac{r \cdot y_1}{\sqrt{x_1^2 + y_1^2}} \\ z_p &= \frac{r \cdot (z_1 - 1)}{\sqrt{x_1^2 + y_1^2}} + 1 \end{aligned} \tag{7.10}$$

7.1.2 Projection Parameters

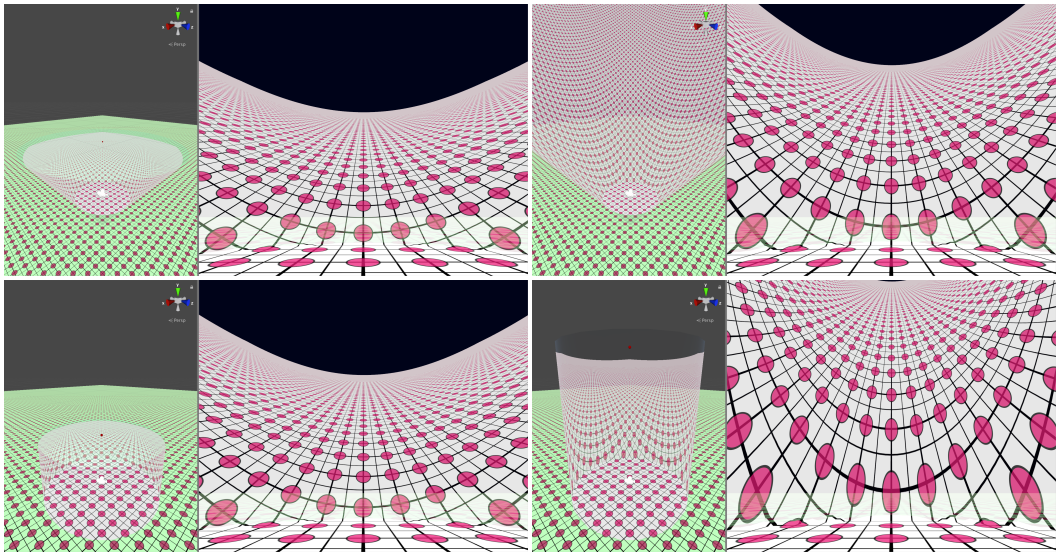


Fig. 7.6: Projection on a cone (top row) and cylinder (bottom), shown their projection origin at the same place as in their equivalent (Figure 5.15) spheres (left column), and set higher for a stretching effect (right column).

Besides the choice between the three projection surfaces explained above and their dimensions, there are a number of other parameters that determine the resulting projection. As discussed in section 5.3.6, for the cone and cylinder the projection origin (x_0, y_0, z_0) shown in Figure 7.3 can be set higher than 1, which will stretch the projected image. Figure 7.6 shows how setting this point very high up (by a factor of 30) this can slightly increase the height of the horizon of the projection in a cone at the expense of moderate stretching just behind transition, while even a moderate increase of this projection origin (by a factor of 2.5) can lead to a much higher horizon in a cylinder with the same intersection, however with more severe stretching.

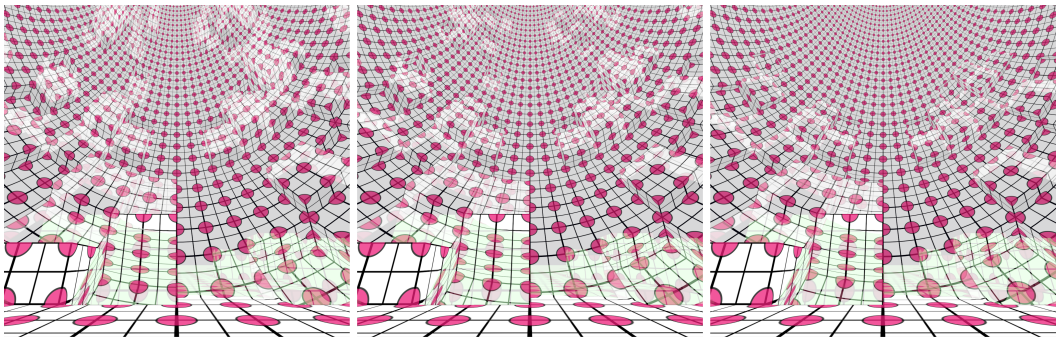


Fig. 7.7: Terrain flattening before inverse stereographic projection to a sphere, seen from the inside. From left to right: no flattening, gradual flattening after the transition to half the original height, and less gradual and complete flattening.

The *terrain flattening* discussed in section 5.3.5 has been implemented with a simple multiplication of the height component for all geometry with a value between zero and one before the transformations and projections described above, which itself is first

multiplied with a circular gradient that starts with zero just outside the intersection and ends with one at a distance presented as another parameter. Together, this "flattening factor" and its "transition distance" allow for flexible control over the appearance of varied terrain as projected, and is shown in exaggerated examples on a synthetic scene with cubes in Figure 7.7 and a view of San Francisco in Figure 7.8.

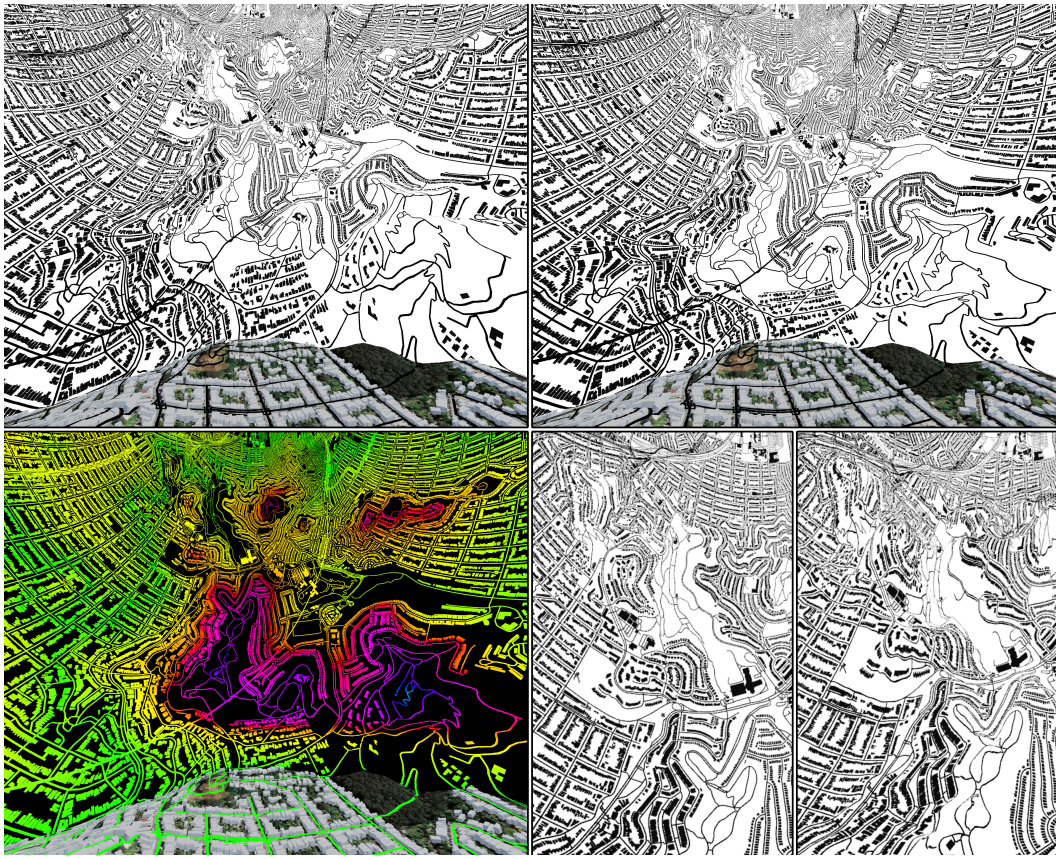


Fig. 7.8: Projection of a hilly city without flattening (top left), and fully flattened before projection (top right). The differences in elevation between the two are shown in a color spectrum (green: no difference; bottom left), and magnified (bottom right).

The above parameters (shape, size, and location of the projection surface; projection stretching; and elevation flattening and transition) dictate every aspect of the deformation. Even among them a clear differentiation exists between the first three (shape, size, location), which can be seen as the *primary* projection parameters, and the latter two (flattening and stretching), which serve a *secondary* role in modifying the appearance of what lands on the surfaces defined through the primary parameters.

The shader can however be extended even further, with Figure 7.8 already showing examples of *tertiary* parameters, such as the choice of culling any original geometry outside the projection surface, projecting only the footprints of buildings and not their whole shape, and conditional coloring of the projection by elevation difference. Figure 7.9 demonstrates further parameters, such as semantic coloration, different shading for geometry outside the projection surface instead of complete culling, and an elevation

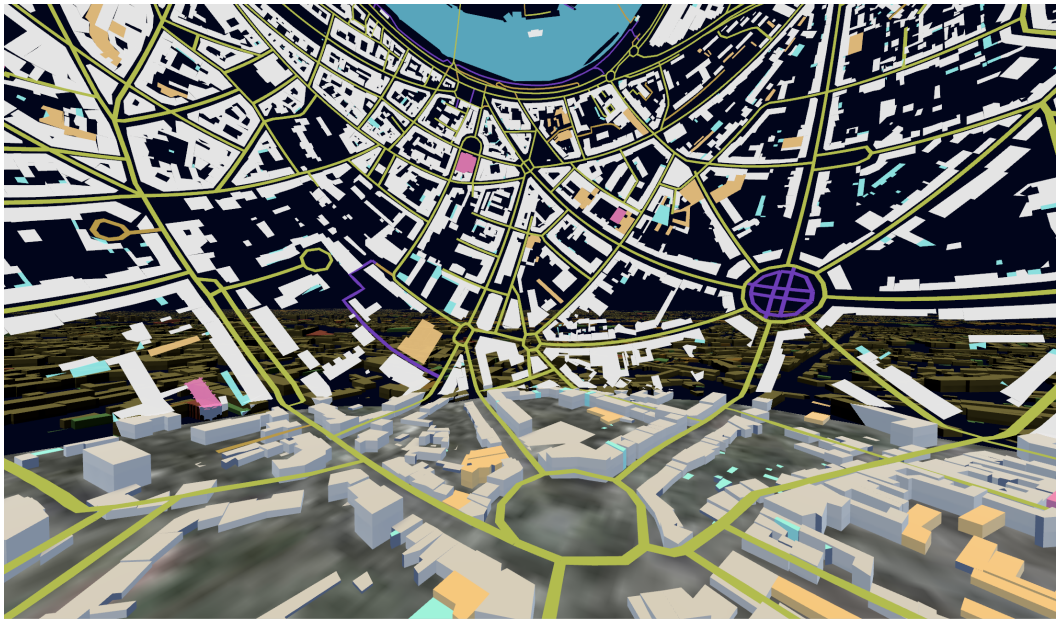


Fig. 7.9: Inverse stereographic projection of a city onto a sphere, with an elevated position of the viewer. Projected are building footprints and street networks, colored by function, as well as the city's river. Buildings outside the sphere are dimmed but otherwise unchanged. Streets are slightly elevated over their original position above the terrain.

of the projected layers over the base terrain *inside* the projection surface — potentially useful for street networks that would otherwise be obscured by 3D buildings.

Figure 7.10 shows a configuration where the full 3D shapes of buildings are projected and not just their footprints, which highlights the final parameter: transparency of the projection. In this example, the camera is positioned at about eye-level in scale of the city, and the original scene is not altered beyond the addition of the transparent projection, in which the color of the streets indicates elevation differences. This approximates a possible application for augmented reality. Figure 10.13 in the annex shows the complete shadergraph as implemented, which includes all parameters discussed here, as well as functionality for an adaptation of Lorenz' geometry-bending approach from [Lor+08].

7.2 Interaction and Navigation

7.2.1 Projection Interaction

All parameters discussed above can of course in theory be adjusted in real time thanks to the dynamic nature of shaders. The distinctions made above between primary, secondary, and tertiary parameters are however useful in deciding which ones to expose to a user for interaction during e. g. exploration, and how.

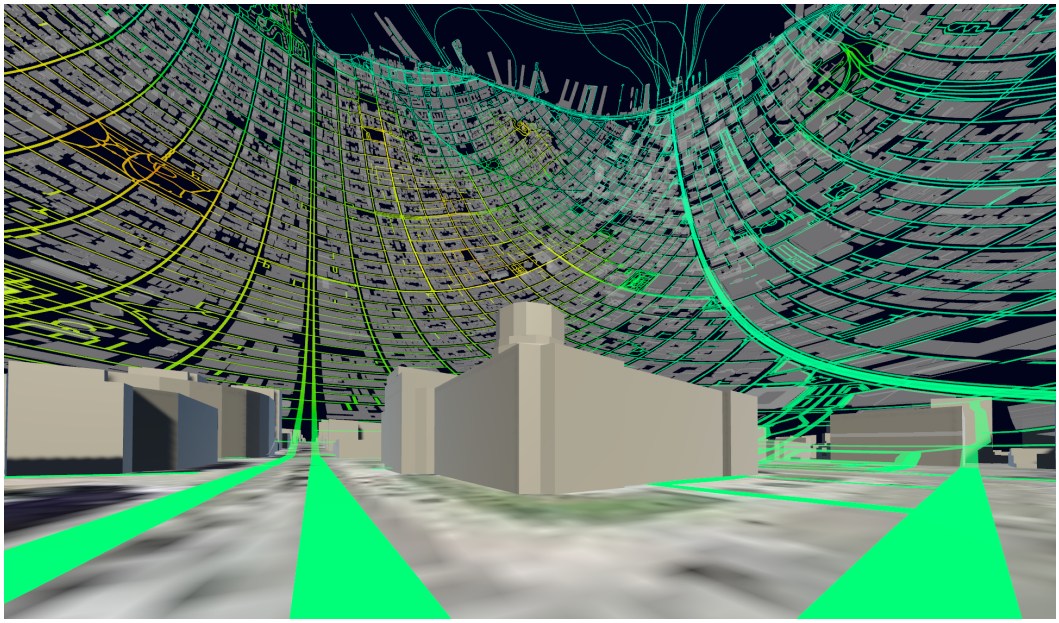


Fig. 7.10: Inverse stereographic projection of a city onto a sphere, seen from a pedestrian perspective. City geometry remains unaffected outside the sphere; streets are colored by elevation, and the 3D shapes of buildings are projected transparently.

The tertiary parameters (coloration, transparency, projection of buildings, etc.) are most likely to be dependent on the use case and available data, and thus most likely to be set up beforehand, with possibly different settings for different objects and layers (e. g. streets, buildings, and terrain). A specialized interface that brings these together is possible, but was not part of this implementation; instead certain parameters such as the culling of the original scene outside of the projection surface were toggled with keyboard shortcuts.

More important for adjustment during traversal of a scene between different positions, heights, and scales are the primary and secondary parameters, which have all been mapped for control with the touchpad of single HTC Vive controller (Figure 6.1).

The secondary parameters, as their name already implies, are accessed with the touchpad while holding a modifier. Terrain flattening can controlled while holding the side buttons: an up-down motion adjusts the flattening factor between zero and one, while a sideways motion adjusts the transition distance. The implementation is such that the initial motion (from the first touch to the next position sample) determines and locks the axis, which makes it easier to adjust only one of these parameters. Lifting the thumb and swiping again is effortless enough that both parameters can still be adjusted quickly if needed. Pressing down on the touchpad while swiping up and down adjusts the height of the projection origin in the cone and cylinder above the default 1, for the effect seen in Figure 7.6.

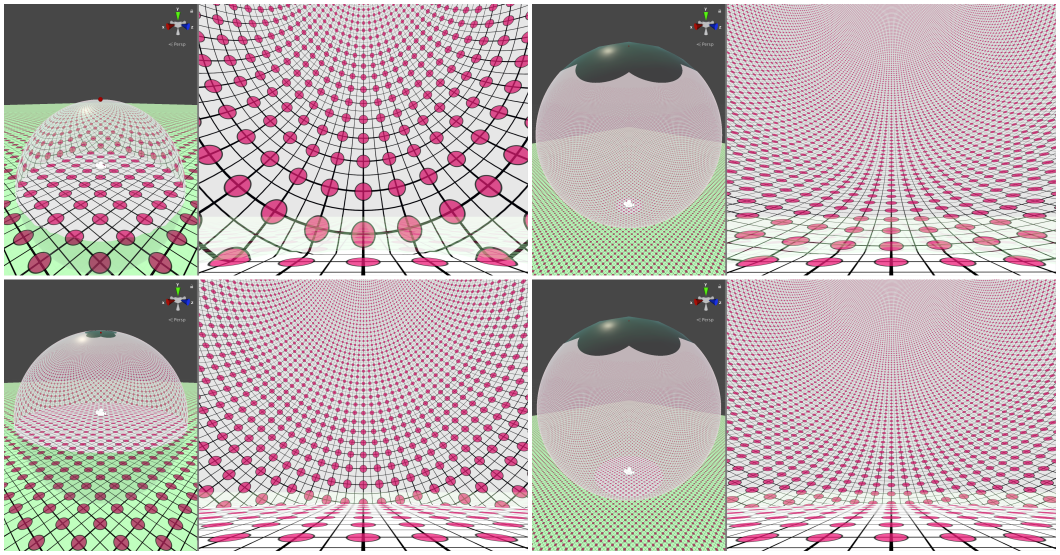


Fig. 7.11: External and internal views of the inverse stereographic projection with different *radii* on the sphere, bisected by the base surface at its equator in the left column. A short *focus view extent* is shown in top row, with a long one below. The *intersection ratio* is smallest in the top right, larger in the bottom right, and at its maximum (50%) in the left column.

This leaves the primary parameters: position, size, and shape of the projection surface. Double-clicking the side buttons toggles between the sphere and its "equivalent" cone (Figure 5.15) — the cylinder is treated as a "special case" of the cone, because the equivalent sphere for it is one that is bisected by the ground plane at its equator, at a right angle, shown in the left column of Figure 7.11. The horizontal position of the sphere is determined by the position of the viewer, as its axis is bound to the virtual camera to guarantee an equivalent view in all horizontal directions. This leaves only the vertical position and the size (or radius) of the sphere (and thereby its equivalent other shapes) as the last two (one-dimensional) parameters which can be adjusted by swiping on the touchpad, this time without pressing modifying buttons. Most trivially, the touchpad can be used to move the sphere up and down with vertical swipes, and change its radius with horizontal swipes on the touchpad. There are however potentially better ways that can create a certain visual consistency in interacting with the projection parameters.

First, the *intersection ratio* (Figure 7.11) can be established as a measure of the bisection of the sphere by an imaginary horizontal plane that passes through the original terrain at the point on the sphere's north-south axis — just below the viewer. Its (constrained) range is $]0, \frac{1}{2}]$ — from the sphere being almost tangential to the terrain, to being cut in half by it, or from appearing as an almost complete sphere to only a dome, as shown in the top row of Figure 7.11. It is the ration between the *intersection height* h shown in Figure 7.12 and the radius. This measure can also be used to calculate the angle α

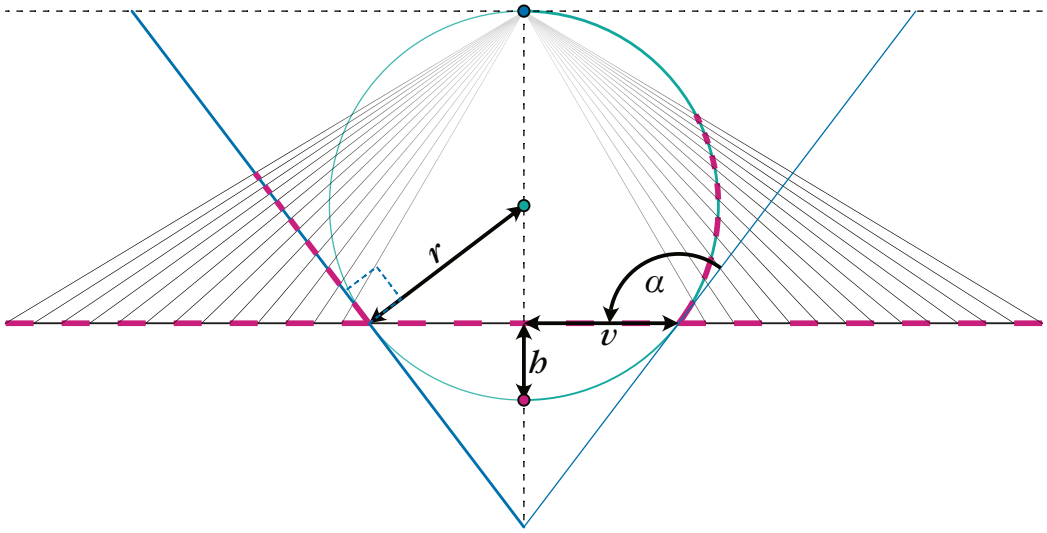


Fig. 7.12: The four linked primary parameters of the inverse stereographic projection: radius (r), projection intersection height (h), intersection angle (α), focus view extent (v).

(Equation 7.11) at which the transition from said horizontal plane to the sphere or cone occurs, and is thus integral to the appearance of the projection close to the transition.

$$\alpha = \frac{\pi}{2} + \arcsin(1 - \frac{h}{r}) \tag{7.11}$$

$$v = \sqrt{r^2 - (r-h)^2} \tag{7.12}$$

As a viewer navigates around uneven terrain, the sphere can thus not only follow the horizontally, but also adjust its vertical position relative to the ground just below the viewer, so as to keep this intersection ratio or angle at a fixed value and preserve the appearance of the transition at all terrain elevations. The same can be done when the user wishes to change the sphere's radius: the vertical position of the sphere will be adjusted to keep the intersection angle/height constant. This change in radius also changes the *focus view extent*, shown as v in Figure 7.12, and which can be expressed as a function of the radius and the intersection height (Equation 7.12).

Apart from the radius, the other two higher-level measures (intersection ratio/angle and focus view extent) describe the *appearance* of the projection: how far from the viewer the transition begins, at which angle the terrain is intersected by the projection, and how large the surrounding sphere is. Any two of these parameters can be freely adjusted by a user, if the third (and the low-level intersection height) is automated to compensate for changes. Figure 7.13 shows an automated intersection angle (by moving the sphere's height) with the focus view extent and radius under the user's control, and Figure 7.14

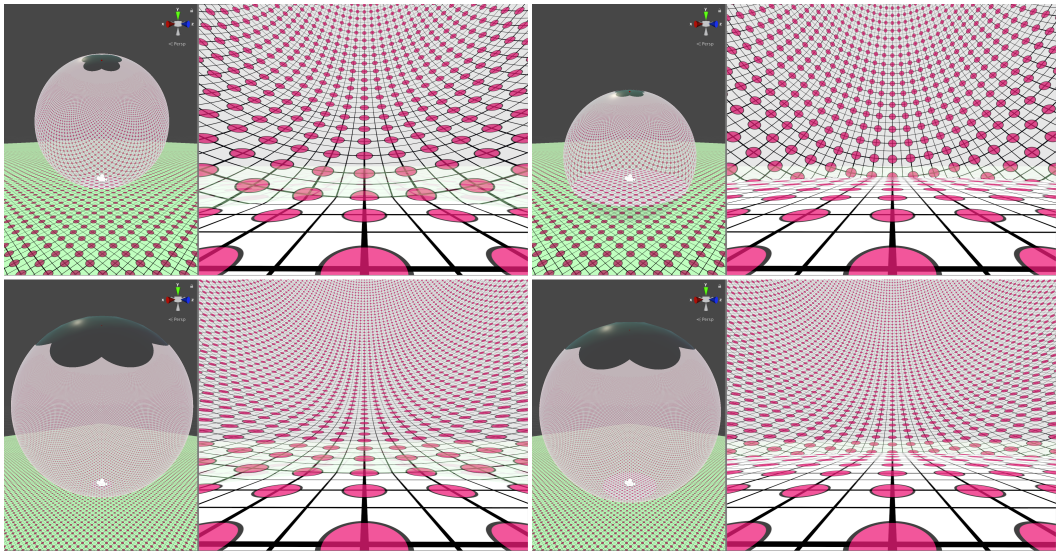


Fig. 7.13: Automatic intersection angle: the same radius within rows, the same focus view extent within columns (increasing to the right and down).

shows an automated radius (and height) with the intersection angle and focus view extent under the user's control.

This automation aims to make control over the projection more intuitive by keeping in place an *apparent*, or high-level parameter, instead of a lower-level parameter which could affect the perceived projection in more than one way. This implementation thus exposes only the three high-level parameters to the user, with a choice between which one of them will be automated, and which two are under control. The chosen two are finally controlled by swiping vertically or horizontally on the touchpad, with a double-click on the side buttons toggling between which of the three parameters should be automated. Table 7.1 shows which axis controls which parameter under which of the three "automation modes."

	extent	**angle**	**radius**
horizontal	angle	extent	extent
vertical	radius	radius	angle

Tab. 7.1: Which high-level projection parameters are accessed by which axis of the touchpad under which condition, i. e. which *other* parameter is being automated (header row).

This leaves the trigger of the HTC Vive controller (Figure 6.1) unused. Being spring-loaded, i. e. always returning to an "off" position when not engaged, it was mapped to a final projection parameter: an interpolation between the projected and original position. Pulling the trigger gradually moves the projected image down along the projection rays, until it is fully depressed and it land on their original position, at which point it also disappears so as to not interfere with the original geometry (Figure 7.15). At the same time, other tertiary parameters, such as coloration of the outside geometry is also interpolated to its projected value, making the trigger a gradual "before and after" tool, which a user

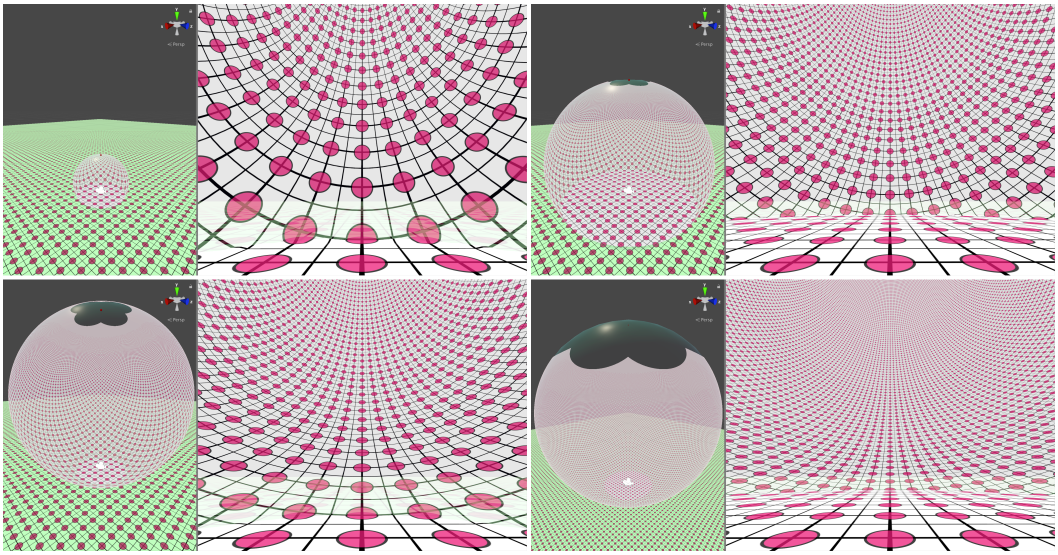


Fig. 7.14: Automatic radius: the same intersection angle within rows (increasing to the right), the same focus view extent within columns (decreasing down).

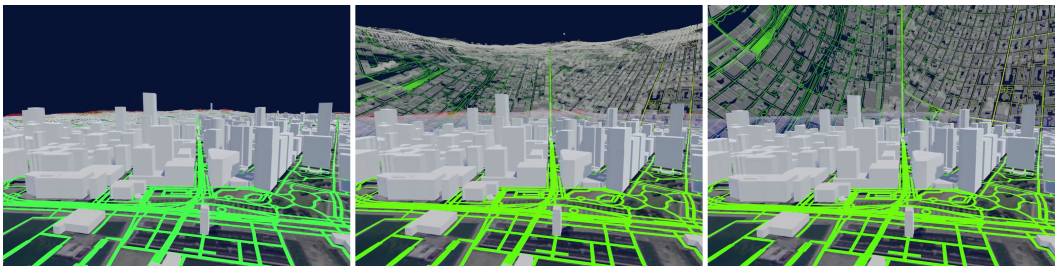


Fig. 7.15: The interpolation (middle) from the original scene (left) to its projection on the sphere (right).

can utilize at any time or speed to perhaps get a more direct overview as to where the projections on the sphere or cone originated from.

7.2.2 Navigation

Having mapped all primary and secondary parameters of the inverse stereographic projection (section 7.1.2) to a single HTC Vive controller, the second one in the set remains to be used for actually navigating the virtual environment. Mostly, navigation in 3D environments is accomplished either gradually or instantaneously, through steering or teleportation, respectively. Specifically for navigation in virtual reality, additional care needs to be taken when choosing or designing methods of locomotion, as movement in virtual reality has been well-known to cause effects such as cybersickness in many users since the "early days" of VR [Kol95], and these effects are still being heavily investigated as HMD technology progresses [CP19]. Generally speaking, steering is a more natural method that maintains a high degree of spatial awareness, buy can easily cause motion sickness, while teleportation usually avoids it because there is much less if any difference

between visual and vestibular stimuli. Teleportation however disrupts *spatial updating* — the "process that automatically keeps track of where relevant surrounding objects are while we locomote, without much cognitive or mental load" [Rie03].

Steering is usually accomplished in an indirect manner, by e. g. tilting a joystick or pressing buttons to change directions and speed or acceleration, however more direct methods have been proposed, such as an approximation of the functionality of a hamster ball [Hur+18a]. For teleportation the user usually needs to locate a target, and point at it to be immediately transport to its position, or nearby it. This can pose a challenge when longer distances or occlusions are involved, similar to the challenges that motivate overview+detail and focus+context visualizations. The projection achieved through the inverse stereographic technique can by its own nature, as defined by its design goals and its implementation described above serve as an aid in both methods, as it envelops the viewer with legible targets to point to or otherwise interact with.

Apart from the locomotion achieved by the HMD tracking the user walking around, a number of methods to move and scale the terrain relative to the viewer — navigation relative to the virtual environment — were implemented. The first one is a similar method as used for navigating the MCV stack (section 6.1.2), here called **Air Grab.** Holding the side button, the terrain becomes horizontally tied to the controller, and follows its horizontal motions. The resulting translation can be scaled up, so that smaller hand movements can be enough to rapidly traverse larger distances. Vertical movements can also be translated to zooming, same as done with the MCV stack.

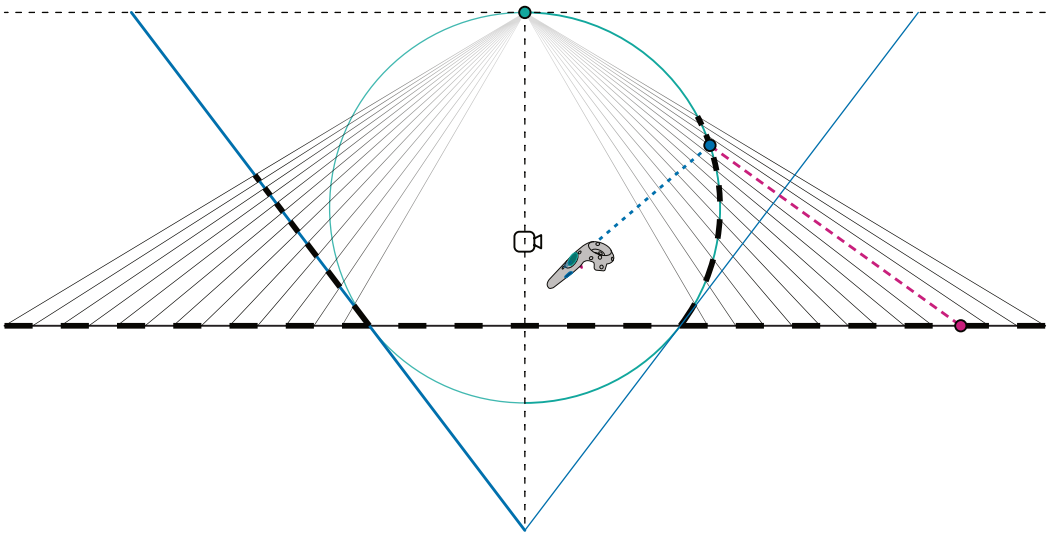


Fig. 7.16: The projection surface's vertical axis is bound to the viewer and moves with the virtual camera horizontally. Ray casts intersecting the surface are projected down to the original terrain using the same stereographic projection method.

All following navigation methods listed below rely on ray casting. Using the trigger, the presence of a thumb on the touchpad or a click of the touchpad, the user can cast a

visible ray from the tip of the Vive controller, which can intersect with the projection surface or the unchanged geometry within it. An intersection at the projection surface can easily be translated to coordinates on the original terrain, using an inverse of the inverse stereographic projection — in effect, the user is pointing to potentially far away targets via a "bend" in this ray, as shown in Figure 7.16. These methods are:

- **Point Grab:** The target is being held by and moved with the pointer, allowing the user to grab the terrain and move it around horizontally. A ray is shown upon placing a finger on the touchpad, and the terrain is grabbed at the target being pointed at when the touchpad is then pressed down.
- **Zoom** and **Rotate:** Pointing initiated as in *Point Grab*, but instead of pressing the touchpad, horizontal swipes across it rotate the terrain around the pointed target, and vertical swipes zoom into or out of the target.
- **Steer:** Pulling the trigger slightly casts a ray. Depressing it further causes the terrain to move towards the viewer from the pointed target's directs. The speed of this steering is dependent on the distance of the target and on how much the trigger is depressed.
- **Fly to:** The ray is initiated as in *Steer* with a slight pull of the trigger. Clicking the touchpad while pointing this way, an automated transportation to the target is executed with smooth accelerations. All input during this transportation is ignored.
- **Teleport:** An instantaneous version of *Fly to*. Initiated the same way as *Fly to* and *Steer*, but executed with clicking the side buttons while pointing. Alternatively, double-clicking the side button toggles between *Fly to* and *Teleport* functionality.

Air Grab, **Zoom** and **Rotate** further implemented a measure of inertia to the movements after their controls are let go for a more natural behavior by adding the perception of a certain "weight" to the virtual environment. More distance can thus be traversed with less movements, while precision is maintained by allowing the user to immediately stop any inertial motion by engaging any of these three methods again.

During all navigation physical or otherwise, the sphere is tied horizontally to the viewer, and the terrain is constantly translated vertically so that the point immediately below the viewer is always at floor level. Alternatively, this could be set a different height between the floor and the user, for e. g. a tabletop-like appearance. Finally, a red arrowhead can be shown permanently hovering above and slightly in front of one of the controllers, which always points to the north of the map to aid in orientation. An example the system as seen through the HTC Vive HMD is show in Figure 7.17.

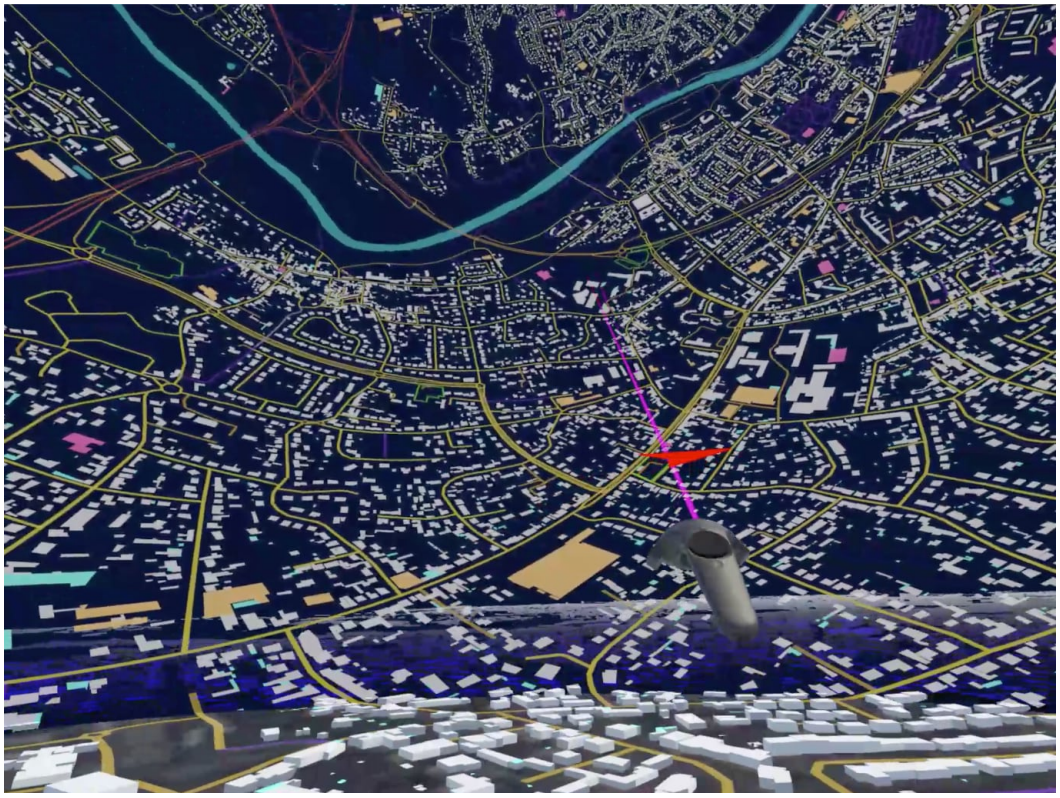


Fig. 7.17: Inverse stereographic projection of a city, showing the controller used for navigation with its compass indicator and casting a ray to the projection for navigation.

7.3 Evaluation, Discussion, and Comparisons

7.3.1 Iteration and Informal Evaluation

As with the immersive MCV methods, urbanists and VR researchers were frequently consulted during the iterative development process. Their input helped inform the choices made regarding parameters for the navigation methods and the controls and constraints over projection parameters, as well as shape aesthetic choices for colors and opacity. Especially the magnitude and the speed of the navigation required very fine adjustments to satisfy the majority of the users.

As the implementation neared a more complete state, a series of exploration tasks were conducted with the system as implemented above. The local city (Nantes) was used as the geographic model, with a satellite map as a texture for the 3D terrain, the road network taken from Mapbox [Map20] and the building shapes from the BDTOPO database [IGN20], configured as shown in Figure 7.17.

Local researchers and professionals in urbanism were given the HMD, which first presented them with a 3D view of their home town from high above (Figure 7.18 left), with



Fig. 7.18: The map exploration system as used for the informal evaluation shown at three different zoom levels; fully zoomed out the whole map can fit inside the focus view extent and is thus not projected.

all navigation disabled. The map being fully zoomed out, it fully fit within the focus view extent of the inverse stereographic projection with its initial parameters, so no projection is seen at that zoom level.

At first, only horizontal navigation methods were enabled and explained, i.ė. no zooming was possible. The users were instructed to then navigate to their own house in this zoomed-out state, as a first test of their understanding of the basic 3D map and basic controls, such as *Air Grab* or *Point Grab*. They were then instructed to navigate to the building they were currently in. At this point, the remaining navigation methods were enabled and explained, which allowed them to zoom in to the current building. Doing so scales the map far beyond the focus view extent and into the context area, which results in something being projected on the sphere's surface for the first time, in a way that hopefully appears understandable to the user.

From this zoomed-in position with the projection now visible, the users were asked to navigate to a certain close-by landmark, preferably now also using the navigation methods on the sphere's surface, without zooming out again. Finally, the commands for interacting with the primary projection parameters were enabled and introduced, enabling the users to adjust the view of the context in a way that allowed them to more easily find a third, much more distant landmark, which concluded the "tutorial" part of this informal evaluation. After that, various tasks were given to the users, such as following a certain road or river, finding and navigating to other landmarks or buildings shown with a certain color, and freely exploring on their own.

All eight participants were seen to quickly master the interaction methods, and some preferences for certain navigation controls over others could be observed. The "grabbing" methods, especially the pointer-based one were strongly favored over the steering methods, which quickly prompted mild cybersickness in some users. Surprisingly, the automated steering in the *Fly to* navigation method was deemed acceptable once the acceleration curves were individually calibrated to the demands of each participant. Shorter travel within around the scale of the focus view extent or below was preferably executed with the *Air Grab* method, most likely as it does not require attention to where

the controller is pointing towards. Longer-distance travel exceeding these distances on the other hand benefited greatly from the context view on the projection surface.

Regarding the novel focus+context view as presented by this implementation of the inverse stereographic projection, most participants — including those with extensive experience with 3D city models — were positively surprised at seeing the city in this way. One professional in urbanism expressed how this view "completely changes [her] way of understanding urban spaces." While some of that sentiment — which multiple users expressed to varying degrees — could perhaps be attributed to just seeing a 3D city model in immersive virtual reality, it could also be possible that the enveloping nature of especially the sphere heightens this feeling of being immersed in the data in a literal way.

7.3.2 Contributions and Effects

There are a number of contributions to visualization in a broad sense that the presented implementation (or the "present system") brings. The following bullet points list the most important of them, the implications and effects of which are discussed in more detail below.

- A novel, omnidirectional focus+context geovisualization method for immersive environments.
- The mathematical extension of the stereographic projection to planes intersecting a sphere at any height.
- The application of the principle behind the inverse stereographic projection to other surfaces with circular symmetry.
- Precise control schemes for easily interacting with parameters of this projection and for quickly navigating inside a virtual environment that uses this projection.
- The whole framework that unifies all of the above, which can be extended with additional focus+context methods, allowing for instant switching between them for greater flexibility or for fair comparisons in controlled studies.

The most basic contributions made while developing this "Virtual Data Sphere" technique are the idea of a novel omnidirectional focus+context view designed specifically for immersive applications that merges a first-person view with a top-down perspective, and the mathematical extension of the stereographic projection into the third dimension that makes the distortions from this multiperspective view conformal, at least in the case when a sphere is used as the projection surface.

Conformal map projections are valued for their preservation of local shapes, but are usually only thought about in the context of projecting a globe to a plane, for viewing as a map on a flat screen or printed out. In this case, the projection surface is a sphere

itself, so unless the viewer is positioned exactly at the center of it, there will always be an additional, perspective distortion in this method. The same however is true for the traditional map projections, and as discussed before in connection with the MCV stack, a certain tilt is usually perceived as just that and not a distortion, with the perceptive system being able to interpret the information as if it was directly facing the viewer [VGB05].

With the sphere being such a simple and common shape, the hope here is that the additional perspective distortion is similarly perceived as a curved tilt — in fact most people made a comparison to the 2010 film *Inception* when first being shown this method; perhaps implying that they see the city as being folded or bent, and not distorted. The precise perception of the projection certainly warrants closer investigation into how viewers actually see it: a merger between different perspectives, a compression, a geometric bend? Does a categorization like this change the perception of e. g. distances? Does it differ drastically between the different projection surfaces, of which only the sphere distorts conformally?

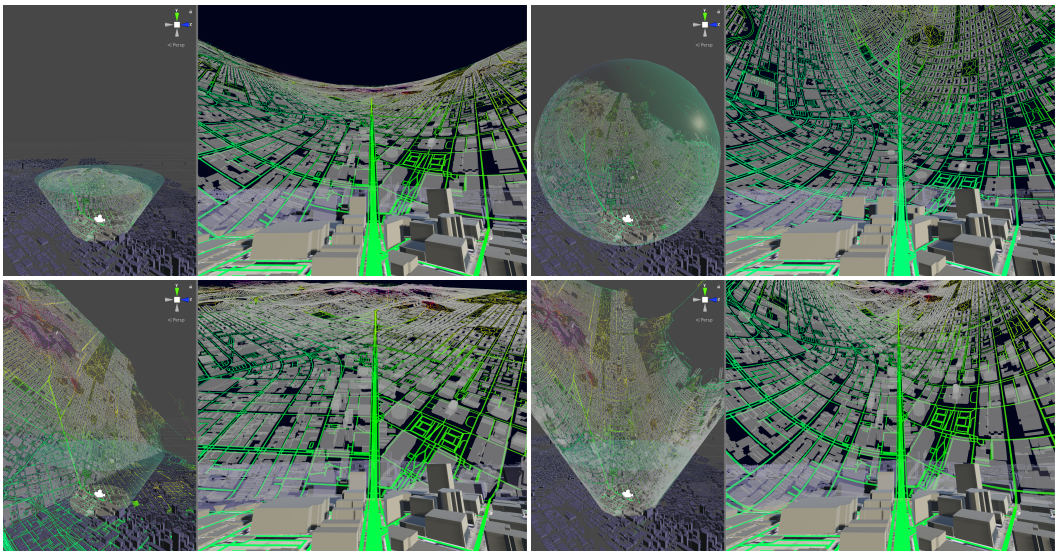


Fig. 7.19: Four deformation methods: Inverse stereographic projections on a cone (top left) and a sphere (top right), and bending along one axis (bottom left) or along a circle (bottom right) that is located at the cone's or sphere's intersection. The cone and sphere are not "equivalent" here, but adjusted for a more comparable inside view.

Ultimately, despite the theoretical benefits of a sphere as a projection surface over a cone or cylinder (conformal distortion, no "horizon," a "simpler" shape), the latter may in fact be preferred in certain cases, such as when projecting more regular street networks/grids, as shown in the top row of Figure 7.19. The downsides of not utilizing the sky above the cone's or cylinder's rim and the much stronger compression towards that new horizon may be outweighed by the perhaps more "regular" appearance of the projection which looks less "bent" — several viewers did express their preference for it, in fact.

The other surfaces are more flexible if their projection origin is detached from its position on their equivalent spheres: The horizon can be raised, and the projection just behind the intersection stretched. If considered as just another parameter in an overall "inverse stereographic projection" system for focus+context, the choice of projection surface just adds to the flexibility of it in providing a "never-before-seen view of the city," as some users have called it.

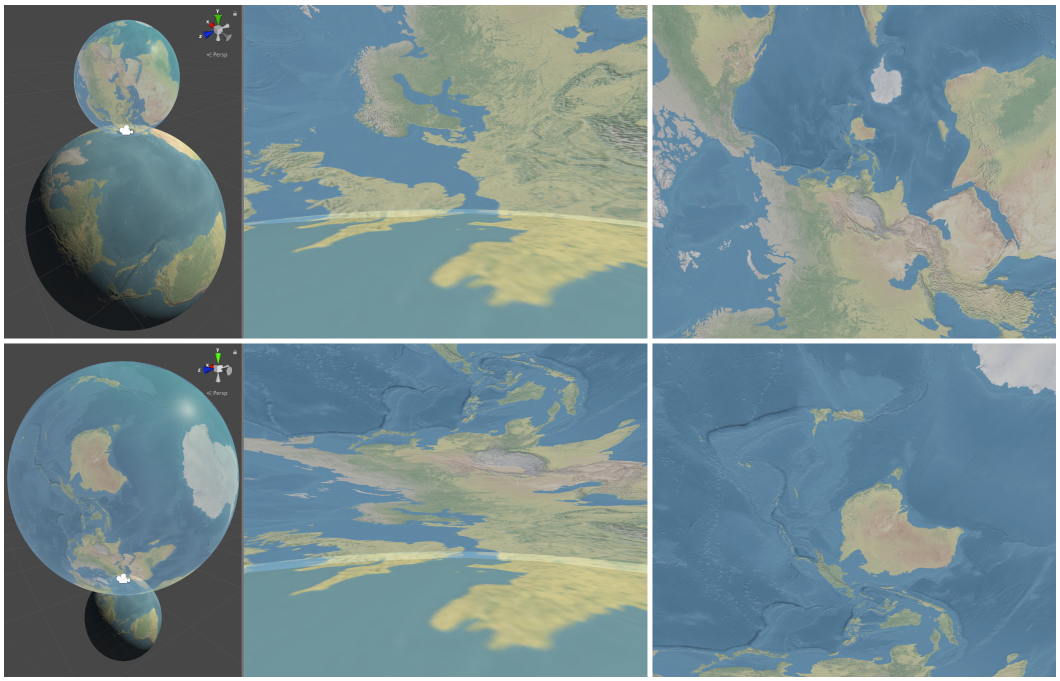


Fig. 7.20: The double stereographic projection described in section 5.3.7, shown with two different sizes of the projection sphere (top vs. bottom), and seen from inside looking toward the horizon (middle) and toward the zenith (right). The projection starts just before the true horizon, with a small overlap visible.

The existence of the horizon necessitates all context being compressed into a smaller space, making an overview perhaps easier, but the same distortion that makes regular perspective projection unreadable close to a "natural" horizon occurs near the rim of the cone or cylinder, resulting in an incomplete overview in the end. The sphere on the other hand, by projecting conformally, does not suffer from this issue, and could in fact legibly project arbitrarily far away, with features merely shrinking toward the zenith. If using the "double stereographic" approach described in section 5.3.7, a whole planetary body can be projected, where the size of (the conformally projected) distant and even antipodal features depends on the size of the projection surface, as shown in Figure 7.20.

The other secondary parameters also greatly contribute to the flexibility of the overall system. Setting how much and how quickly after the intersection the terrain is flattened before projection allows the user to balance the effects of exaggerated elevations and the resulting structure-from-motion effect with a less distorted, or even fully flat map-

like appearance, all by using simple real-time controls and without having to leave the immersive environment.

The tertiary parameters, dictating how (and if) objects appear inside, outside, and on the projection surface can crate a wide variety of configurations through differentiation in these regards. Combined with the often strongly apparent "fold" at the intersection it can be made very clear what is the original scenery, and what is the added projection. Conversely, choosing a much shallower intersection angle and making the projection have the same appearance leads to much smooth transition from focus to context views, if so desired. An even shallower intersection angle, achieved by e. g. a sphere with a very large radius, in combination with a selective projection of e. g. only the street networks can create a kind of multi-layer visualization in a similar spirit and motivation as the MCV stack — separating layers of geospatial data vertically, while maintaining a measure of visual continuity.

All primary and secondary projection parameters were mapped directly to the interface afforded by one of the HTC Vive controllers for quick real-time access. The other controller was fully utilized for navigating around the virtual environment. Being able to use both the original scenery inside the focus view as well as the context projected on the surfaces for raycast-based navigation like targeted teleportation or "grabbing" allows for both fast travel over very large areas and very precise local movements using the same methods without switching modes — the compression coming from the projection creates a smooth transition between the two by itself. With the headset as yet another input method — controlling the view direction — the system allows for a much more fluid, fast, *and* precise interaction with maps than is common or perhaps even possible with traditional desktop environments or smartphones. Even disregarding its focus+context aspect, the implemented system could count as a contribution to highlighting the possible advantages of immersive environments have for geovisualization and navigation over previous methods.

This same flexibility is however complicating formal evaluations of the system. section 6.2 is an example of the difficulty in arriving at and setting up a study with any ecologically validity that involves visualizing multiple layers of geospatial urban data, and there the system itself was rigid. In the first qualitative evaluations discussed above, the tasks were only concerned with finding and navigating towards certain objects, not with the users e. g. attaining insights into the nature of both focus and context areas at the same time. Same as the study on comparing MCVs resulted in deeper insights into user behavior by finely measuring performance and tracking eye movements, correlated with demographic information, a rather simple study involving only navigation or exploration tasks could reveal perhaps patterns in what projection surface is preferred by which users, with which parameters, and how they navigate it.

The framework created for implementing the different projection surfaces itself is also a major contribution. It unifies all projection surfaces under the same interaction schemes and allows instant switching between them, giving the system a flexibility that was found to be desired in the MCV comparison — users can very quickly adapt the projection to their exact needs and preferences. Furthermore, it can be extended to include other focus+context methods while retaining the same interaction schemes, for either comparing them as directly and "fairly" as possible in studies, or for giving users even more options in an all-in-one system.

The bottom left image of Figure 7.19 shows how the geometry-bending approach, similar to the one by Lorenz et al. [Lor+08] was implemented in this system (including a variable bend radius for a smooth transition), using the cone as a "guide" for its parameters, similar to how the cone follows its "equivalent" sphere. Similar to their implementation, the direction of the bending axis can follow the user's gaze, or it can be controlled explicitly. The bottom right image shows an omnidirectional version of this bending method, in which the bending axis is a circle rather than a line, as introduced by Vallance and Calder [VC01]. Each one of these systems, and potentially more, can be quickly toggled and adjusted using the same input methods.

7.3.3 Comparison to Existing Focus + Context Methods

The inverse stereographic projection could be seen as a comparable approach to the complex logarithmic panorama introduced by Böttger et al. [Böt+08], in that both result in a smooth and conformal compression from focus to context, and yet this is already where the similarities between it and this system end. At a high level, the complex logarithmic panorama approach results in one static image, intended for printing or viewing on a screen, while the present system is implemented specifically for immersive environments and real-time control over the deformation. The same is true for other purely 2D approaches, such as those by Pindat et al. [Pin+12] or Zhao et al. [Zha+12] — while they certain can displayed as images floating in a virtual space, they would not take any advantage of it other than showing perhaps larger versions of it.

The multiperspective views investigated by Wu and Popescu [WP16], while definitely applicable for interactive and immersive use, are mainly intended for local disocclusion, or selectively applying deformations to reveal hidden perspectives, while the present system globally shows the context around the viewer's position (which is defined as the focus area) in an omnidirectional way.

This only leaves the geometry-bending approaches by Vallance and Calder [VC01] and Lorenz et al. [Lor+08], and their subsequent enhancements by Pasewaldt et al. [Pas+14] — these are in fact so similar, that they have been implemented as variations in the

presented system, at least the simple versions that only bend along one axis or circle, shown in Figure 7.19.

Lorenz et al. contrast their system, which only bends the context along one line to the circular one by Vallance and Calder by saying how their deformation has a "planar and hence, clear and undistorted view of distant regions in the background," unlike the circular one. As previously discussed, and as Figure 7.19 shows, even their "planar" deformation quickly results in non-conformal compression towards the horizon, which makes their claim valid only up to a certain distance. It is however true that the view after the linear bend is not deformed *per se*, only tilted, whereas all circular methods (bends and projections) introduce a deformation that makes originally straight lines appear bent.

Figure 7.19 shows the differences in this deformation between each method: lines appear strongly bent from the beginning in the circular bending method, and appear more straight towards the horizon. Stereographic projection to a cone similarly leads to lesser bent-looking lines towards its horizon, but already starts off less deformed shortly behind the intersection. The inverse happens in the sphere: it begins very similarly as the cone (being tangential to it at the intersection), but increases the apparent curvature toward the zenith.

These deformations and distortions are part of a trade-off. Lorenz et al. hint at the purpose of their linear bending method for a pedestrian view:[2] it answers the question "Where am I going to?", asked by users who "want to look ahead the path along they are currently walking." Even though it could be assumed that a path lies straight ahead and only that path needs to be bent upwards, Lorenz et al. implemented a more flexible solution in which the bending axis remains perpendicular to the camera direction. Instead of not showing any context at all if the user looks sideways or behind, the bend follows the camera, and thus changes the deformation with every camera rotation. This could be acceptable in desktop applications, but may cause distraction at best, or disorientation and cybersickness at worst in an immersive environment, where the camera is tied to the user's head movement, which is never fully still. The circular deformations, in contrast, provide a rotation-independent, panoramic view of all context, at the price of a deformation of straight lines.

How much the appearance of straight lines is distorted stands in a trade-off with how much of the context is to be shown legibly, and where. The circular version of the geometry-bending technique as introduced by Vallance and Calder shows exactly the same amount of context as the same configuration of the straight bending method by

[2] As opposed to a bird's-eye view, in which the viewer looks (almost) straight down on a scene from a greater height, and the deformation bends the geometry *away*, or down, to create a view of a horizon where there was none seen before.

Lorenz et al. if viewed straight ahead, but bends straight lines strongly to the left and right of the camera's direction. A stereographic projection on the equivalent cone lowers the horizon, which brings with it a faster onset of compression, but straight lines do appear much less bent, as seen when comparing the top left and bottom right images in Figure 7.19. Finally, the sphere is the only method in which there is no horizon, i. e. much more of the context remains legible, and it is not being distorted beyond recognition as in the other three methods towards their horizons, rather just shrunk conformally towards its zenith. The trade-off here is a warping of straight lines similar to that of the circular bending near the intersection, but which also does not straighten out with distance, unlike the cone or circular bend.

In these regards, the inverse stereographic projection onto the sphere has a clear advantage over the linear bending technique in the amount of context it can visualize, at the price of straight line deformation. This deformation, to different extents, is the expense all circular methods must incur for providing a static, panoramic view that does not change with a rotation of the view, which may be necessary in an immersive environment. The projection on a cone as such a panoramic method is a compromise between the amount of context rendered legible and line deformation, as compared to the circular bending method.

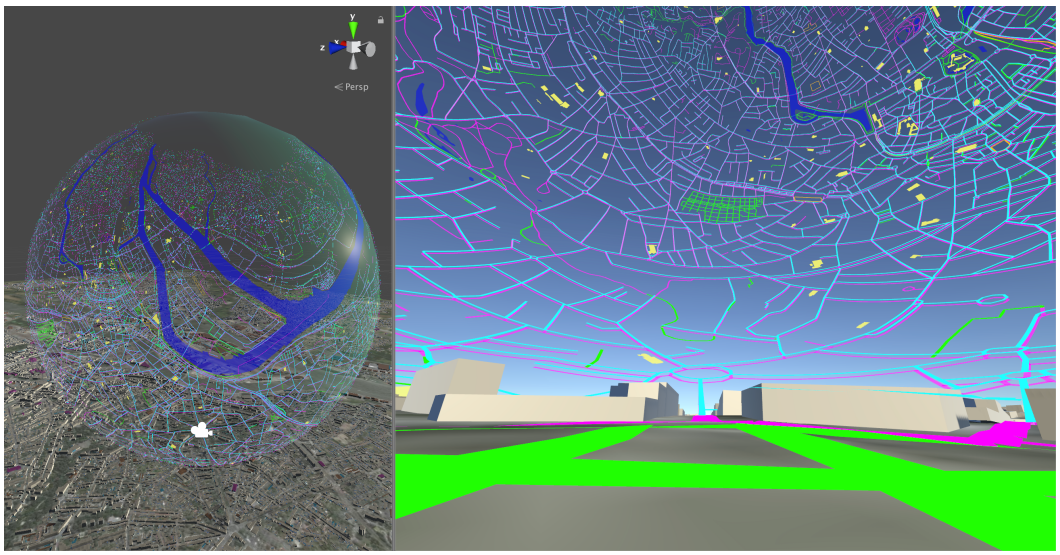


Fig. 7.21: Inverse stereographic projection of only the streets, rivers, and landmarks, seen from a pedestrian view. The cyan streets have been slightly vertically displaced relative to the magenta ones to show overlap.

The later, more complex bending approach by Pasewaldt et al. [Pas+14] does not rely on a single bending axis with a smooth transition zone, but rather relies on a parametric B-spline curve which the deformed context follows. While allowing more control over the compression of context and e. g. the creation of a horizon similar to the original one, the reach of the context is still limited in comparison to the inverse stereographic projection on a sphere. The same limitations and considerations as discussed above with

the simple bending schemes apply, with the addition of a much more complex control scheme that would be required to adjust the parameters of the B-spline curve, as opposed to only needing to change the focus view extent, the angle of the bend, and the length of the smooth transition zone. Computing these transformations is more complex as well, which could adversely impact the frame rates critical to an immersive experience. As can be seen in Figure 7.9, Figure 7.21, and Figure 7.22, using translucency in the projection can reveal the original horizon without adding more complex deformation schemes, if it is so desired.

This latest point also plays to another difference between the previous systems based on bending methods ([VC01; Lor+08; Pas+14]) and the here presented system. As it was developed initially as a multi-layer solution (section 5.1), the projected or bent context appears as just another layer on top of the original environment. The tertiary parameters determine which objects get projected and how, and if the original environment outside of the projection surface gets altered, typically for reducing the visual clutter two identically-looking, overlapping views would cause. The examples listed above in which the original horizon is still visible showcase balances that can be achieved between the original and projected views. In contrast, the previous systems do not keep an unmodified view that can be blended in, for either a view of the horizon, or a multilayer visualization. Instead these fully opaque techniques have been observed to lead to problems with interpretation as to what is original and what is deformed, especially the ones based on B-splines: "the hybrid perspective [...] could be interpreted as [a] hilly landscape." This could be further compounded by the smooth transitions, and especially the ambiguous shape of the free-form B-spline method.

The authors resorted to implementing various cartographic stylizations or different levels-of-abstraction for each zone to "prevent such confusion" — the same could of course be done as additional tertiary parameters for the present system, but the clear intersections with the projection surfaces and the visibility of the original environment in the background should already make the difference between the projection and the focus area quite obvious. This method also allows to only project the context that is actually desired and leave the rest fully intact, e. g. just the road network and some landmarks for navigation purposes, as shown in Figure 7.21. In this, the present system calls back to the MCV stack and the vertical separation of layers, made even more coherent by the intersection that connects them. Figure 7.21, and especially Figure 7.22 show how another tertiary parameter can be used to vertically separate individual layers, such as near-identical road networks to highlight differences between them. This is achieved by elevating the entire projection surface for a layer, along with the original part inside the focus area to avoid discontinuities.

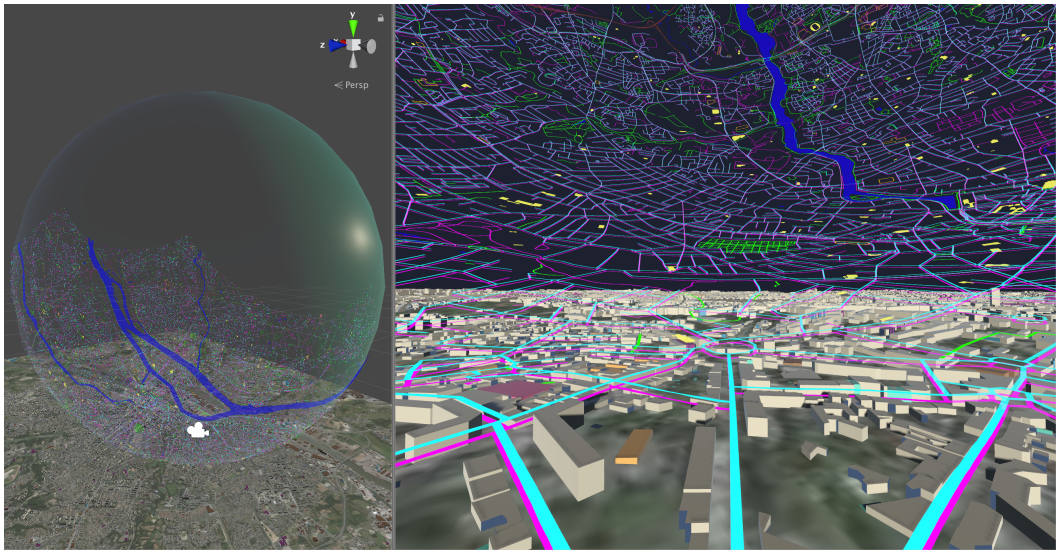


Fig. 7.22: A similar projection as in Figure 7.21, with a larger sphere and seen from an elevated position. This configuration showcases the multilayer capability of the system.

In summary, the presented system appears to be more flexible, more interactive, and more immersive than the most comparable techniques for the same use cases, and enables certain others that would not be practical. The system devised to facilitate and control the inverse stereographic projection can also accommodate those previous methods with the same controls, switchable in real time, resulting in even greater flexibility and choice for users.

7.4 Domains for Applications

The conception of the inverse stereographic projection was motivated in the context of general focus+context visualization for urban environments (section 5.1). During its development it quickly became apparent how it could be applied to uses that reach beyond creating views comparable to the bent geometry achievable by comparable systems, like those by Lorenz et al. [Lor+08] and Pasewaldt et al. [Pas+14]. Going too far into details for each of these would exceed the scope of this thesis, but they provide ample avenues for future work. Some examples are touched upon below.

7.4.1 Urban Development

In the originally intended domain — urbanism — the integrated system based around the inverse stereographic projection could already contribute beyond the capabilities of comparable systems. The fully immersive nature of its implementation in virtual reality makes the system already more engaging than any representation on a desktop screen, even without the focus+context visualization. With most people currently having

experienced VR so far only as a novelty or in video games, if at all, a high proclivity for playful exploration of the urban scenes presented during the informal evaluation was observed in the participants of the informal evaluations, along with strong engagement with the projection. This sense of immersion and engagement is likely to be heightened by the geometry of the projection surfaces, which actually surround or even envelop the viewer.

As discussed before, this engagement is especially desirable during the phases in urban development projects when public consultation, communication, and feedback is sought. An easily navigable 3D model of the city in VR would by itself be a great asset for that purpose, and adding the capabilities of the inverse stereographic projection could further emphasize e. g. the context of a certain area, allow faster and more interesting navigation between multiple distant locations, and allow for direct visual comparisons between them, while at the same time showing their connections through roads, rivers, and railways.

For the urban planner working on e. g. new routes for a public transportation network a possible use-case scenario involving urban connectivity could look like this:

1. An existing transportation hub (e. g. a tram terminal) is viewed from a pedestrian perspective.
2. To gain a broader picture, the first-person camera is lifted above the rooftop level, but the view of the context is still obscured by buildings, especially that of the streets. Translating too much to adjust for a better view can cause cybersickness, and moving away too far/too high will lose sight of the focus area.
3. To see the tram network immediately around the focus area, the planner selects that layer for projection on a cone, where straight lines are better preserved. For better orientation, the 3D shapes of landmarks are added to the projection.
4. Looking at the projection, similar hubs nearby can be found very quickly and instantly navigated to using the raycasting teleportation methods.
5. To see the whole transportation network, extending beyond the city borders, the planner switches to the sphere as a projection surface and adjusts it accordingly.

This possible scenario spans scales from local 3D models of bus and tram stations to the city-wide (and possibly inter-city) transportation network. More local use-cases that could benefit from the kind of focus+context visualizations afforded by this system include visualizations of city zoning and regulatory 3D envelopes, where again the focus could be a single building or city block, and the slightly curved context shows the surrounding land use. Existing buildings could be colored by usage, type, energy utilization, etc., similar to what was demonstrated with the *LakeSIM Integrated Design Tool* [Ber+15], for placing a project in its context, or to show more distant areas, easily accessible, for comparison among themselves and to the focus area.

Using augmented reality, a planner could bring these kinds of context seamlessly into view during on-site visits. What remains for future work (aside from implementing these features in a directly usable way) is to study the impact the kind of immersion this focus+context method, utilizing immersive technologies, has on the *urban* immersion that is so integral to the work of city planners.

7.4.2 Augmented Navigation



Fig. 7.23: The current state of regular, "aerial"/overview (top) and AR (bottom) navigation assistance in consumer products. From Apple's June 2021 WWDC Keynote [App21].

The latest point alludes to a domain adjacent to urban design, but not directly a focus of this thesis: using augmented reality for urban navigation. Most recently adopted by Apple for its Maps App in iOS 15, similar "mainstream" AR navigation aids (by Mapbox, Google, etc.) are currently limited to overlaying arrows over an urban environment, indicating the next turn to take, while their regular navigation views are at best simplified 3D maps seen from an overhead perspective (Figure 7.23).

As far back as 2009, Kim and Dey simulated a system similar in its effect to the some of the views the inverse stereographic projection can provide in a configuration like e. g. that shown in Figure 7.21 for "windshield displays" [KD09]. Using a simple deformation akin

to the one implemented by Lorenz et al. [Lor+08], a stylized street network is overlaid over the real streets, and bent upwards after a certain distance to reveal at least the part of the route that lies in the current heading.

The seemingly independently developed focus+context system by Möser et al. [Mös+08] (Figure 5.6) explicitly states "wayfinding and navigation" as its aims in the title of its publication, but shares its limitations of deforming the whole environment and being mainly intended for viewing on a desktop screen with the implementations by Lorenz et al. and Pasewaldt et al. [Pas+14], whereas the sparse nature of street networks all but invites a separate deformation and transparent overlay of only that layer, as done by Kim and Dey and demonstrated here. Similarly, Veas et al. [Vea+12] used augmented reality in combination with the geometry-bending method to reveal individual buildings/objects or sparse layers, like contour lines.

Visualizing the street network in augmented reality from a first-person, pedestrian perspective using the inverse stereographic projection for the street network with the route highlighted, and possibly certain added landmarks thus seems like a natural fit for navigation purposes, whether from the pedestrian or automotive perspective, or for the elevated view shown at the top of Figure 7.23. All considerations above for the visualization of the context apply. In summary: especially the projection to a sphere can seamlessly connect the street a user currently on (the focus) to their destination in the projection, which can smoothly and legibly compress the entire route, while being unambiguous about directions — users would not have to mentally establish a connection between what is shown on a map with their current position and heading; the projection does this explicitly.

As an additional feature, distances from one's current (physical or virtual) location can be more easily judged or compared, as they directly correspond to a target's vertical position on the projection surface — radial distances from the point below the viewer are mapped as differences in heights in the projection. A circular grid, centered on the viewer, could be added to the projection to make distance comparisons even more direct. Instead of a plain circular grid, *isochrones* [OMS00] could be overlaid on top of the projection to display how far any given transport method could take the user from the current position in a given time. Finally, a combination of these aids could be possible by distorting the projection so that the isochrones fit concentric circles, as has been demonstrated by Hong et al. [Hon+17] for their "interactive distance cartograms to support urban travelers." This way, the height of the now distorted targets on the projection surface would directly indicate their *temporal* distances, and a regular circular grid could now show travel time instead of physical distances.

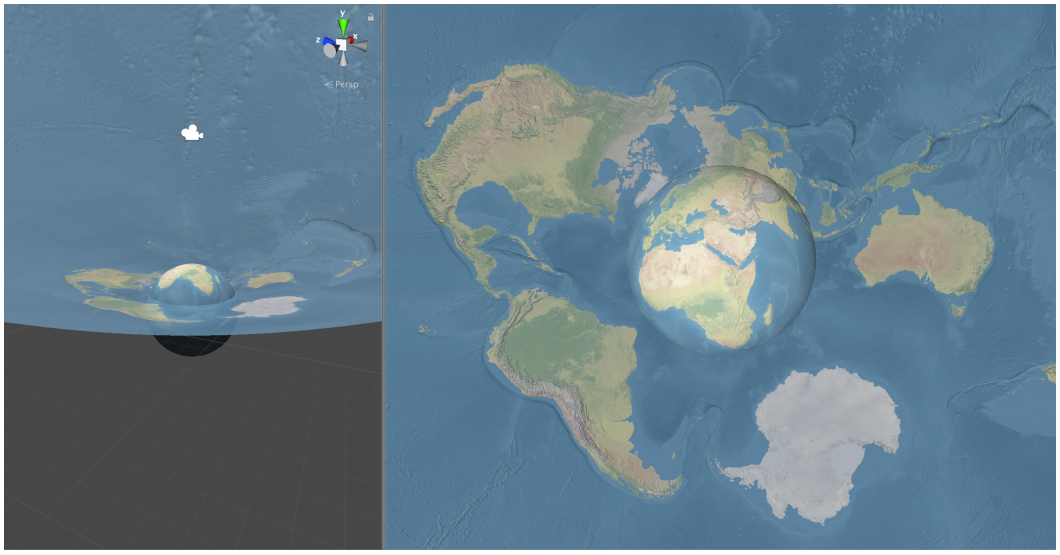


Fig. 7.24: The double stereographic projection (section 5.3.7, Figure 5.17) applied to a globe of the Earth. The intersection occurs exactly at the horizon: the projection shows everything behind it, conformally distorted.

7.4.3 Planetary Visualization

As already stated before and shown in Figure 7.20, the *double* stereographic projection method (section 5.3.7, Figure 5.17) is capable of projecting a complete planetary surface. In geovisualization, the full globe can often be required to show global trajectories and connections [Yan+18b], global atmospheric and oceanic patterns among many other aspects of the *Digital Earth* [Goo+12; MAS15; GGA20], or for exploration of multi-scale environments in VR [HC19]. Even regular, consumer-oriented online and mobile mapping services are gradually introducing a globe view when zoomed out instead of showing the highly misleading Web Mercator projection [Bat+14; LOD18], with Apple most recently announcing a globe view that seamlessly transitions into local street maps, as opposed to previously having to switch between 3D and planar projection modes [App21].

While being the only true representation of the Earth, globes are of course limited to showing at most one hemisphere, and that only if viewed from an infinite distance, which of course would lose any perception of depth. The double stereographic projection can reveal what is behind the horizon without discontinuities or non-conformal distortions (Figure 7.24), and can act as an interaction and navigation target similarly to how it is described in section 7.2.2.

Beyond the Earth, all other planets and dwarf planets are by the International Astronomical Union's definition possessing a roughly spherical shape, as do many more massive moons, making them similarly well-suited targets for application of the double stereographic projection where needed for the purposes of planetary astronomy [Lan+21]. In that field, globe visualization [Bla+18] is already advancing rapidly with the development

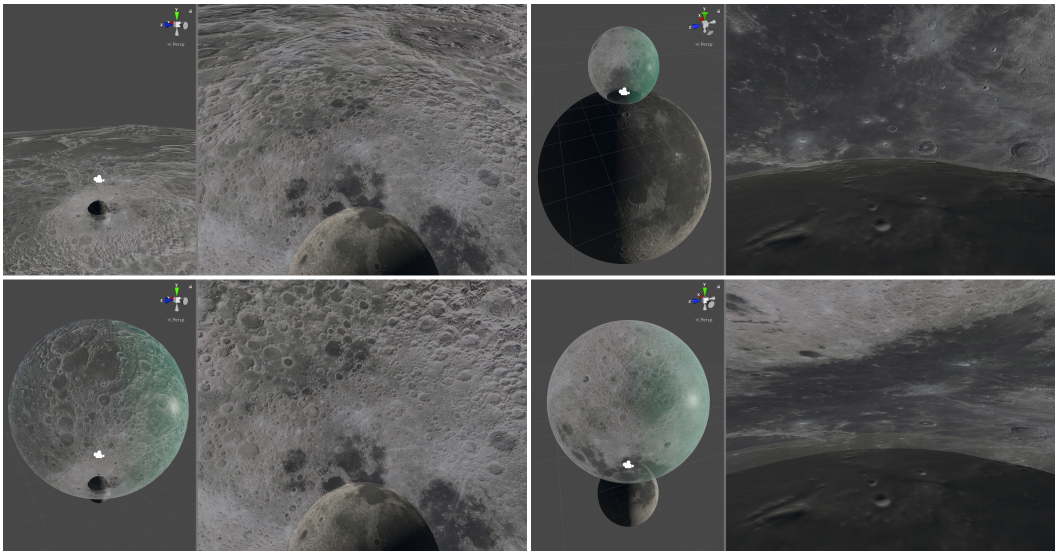


Fig. 7.25: The double stereographic projection (section 5.3.7, Figure 5.17) applied to the Moon. The top left shows the intermediary step of projecting the original surface first to an intersecting plane, before projecting it again to the viewing surface. The other images show different configurations, including intersections happen before the true horizon.

of tools like OpenSpace platform [BHY18; Cos+21], and browsing these globes at different scales could be aided by the double stereographic projection in the same way that it can arguably accelerate navigation within urban areas. The conformal nature of this projection can be seen in Figure 7.25, where the shapes of lunar craters may remain recognizable as near-perfect circles, even when the intersection starts at a shallow angle, seen in that figure's bottom right image.

Focus+context applications in this domain could include improving the visualization of orbital launches, approaches, and landings for planning and public communication [BHY18], and aid in such maneuvers in realistic spaceflight simulators and (educational) video games [RR13; Mal19] like it could for urban navigation (section 7.4.2), or perhaps even assist in the same capacity in actual spaceflight.

7.5 Conclusion



Fig. 7.26: A view of the Earth rising behind the Moon, overlooking Mare Smythii. Taken during the Apollo 10 mission on May 23, 1969. NASA index AS10-27-3896.

Upon reflecting on the pictures taken during the Apollo 10 mission around the Moon [Jac+09], its commander Thomas Stafford said the following about the Earthrise they captured, shown in Figure 7.26:

> This is my favorite photo from Apollo 10. [...] To see our home planet from this view was absolutely awesome. It was nearly breathtaking. I was reminded of Socrates' saying in 399 BCE, before flight above the ground was theorized: "A man must rise above the Earth to the top of the atmosphere and beyond, and only thus will he fully understand the world in which he lives."

The original quote, as recorded by Socrates' pupil Plato in his last dialogue *Phaedo* and translated by Fowler in 1914 [Pla+23] reads rather less lofty in its context, but is still prescient in its own way:

> I believe that the earth is very large and that we who dwell between the pillars of Hercules and the river Phasis live in a small part of it about the sea, like ants or frogs about a pond, and that many other people live in many other such regions. For I believe there are in all directions on the earth many hollows of very various forms and sizes, into which the water and mist and air have run together; but the earth itself is pure and is situated in the pure heaven in which the stars are, the heaven which those who discourse about such matters call the ether; the water, mist and air are the sediment of this and flow together into the hollows of the earth. Now we do not perceive that we live in the hollows, but think we live on the upper surface of the earth, just as if someone who lives in the depth of the ocean should think he lived on the surface of the sea, and, seeing the sun and the stars through the water, should think the sea was the sky, and should, by reason of sluggishness or feebleness, never have reached the surface of the sea, and should never have seen, by rising and lifting his head out of the sea into our upper world, and should never have heard from anyone who had seen, how much purer and fairer it is than the world he lived in.
>
> I believe this is just the case with us; for we dwell in a hollow of the earth and think we dwell on its upper surface; and the air we call the heaven, and think that is the heaven in which the stars move. But the fact is the same, that by reason of feebleness and sluggishness, we are unable to attain to the upper surface of the air; *for if anyone should come to the top of the air or should get wings and fly up, he could lift his head above it and see, as fishes lift their heads out of the water and see the things in our world, so he would see things in that upper world; and, if his nature were strong enough to bear the sight, he would recognize that that is the real heaven.*

This line of thinking predicts in many ways humanity's advance to "the heaven" during the Space Race two and a quarter millennia later, often seen as culminating in the Apollo Moon landings. U.S. President John F. Kennedy's 1963 address that propelled the United States' lunar program is most known for the phrase "We choose to go to the Moon [...], not because it is easy but because it is hard" [Ken63], which can be seen as directly challenging the "feebleness and sluggishness" Socrates attributes to us in our inability to fully perceive heaven and earth from our limited perspective.

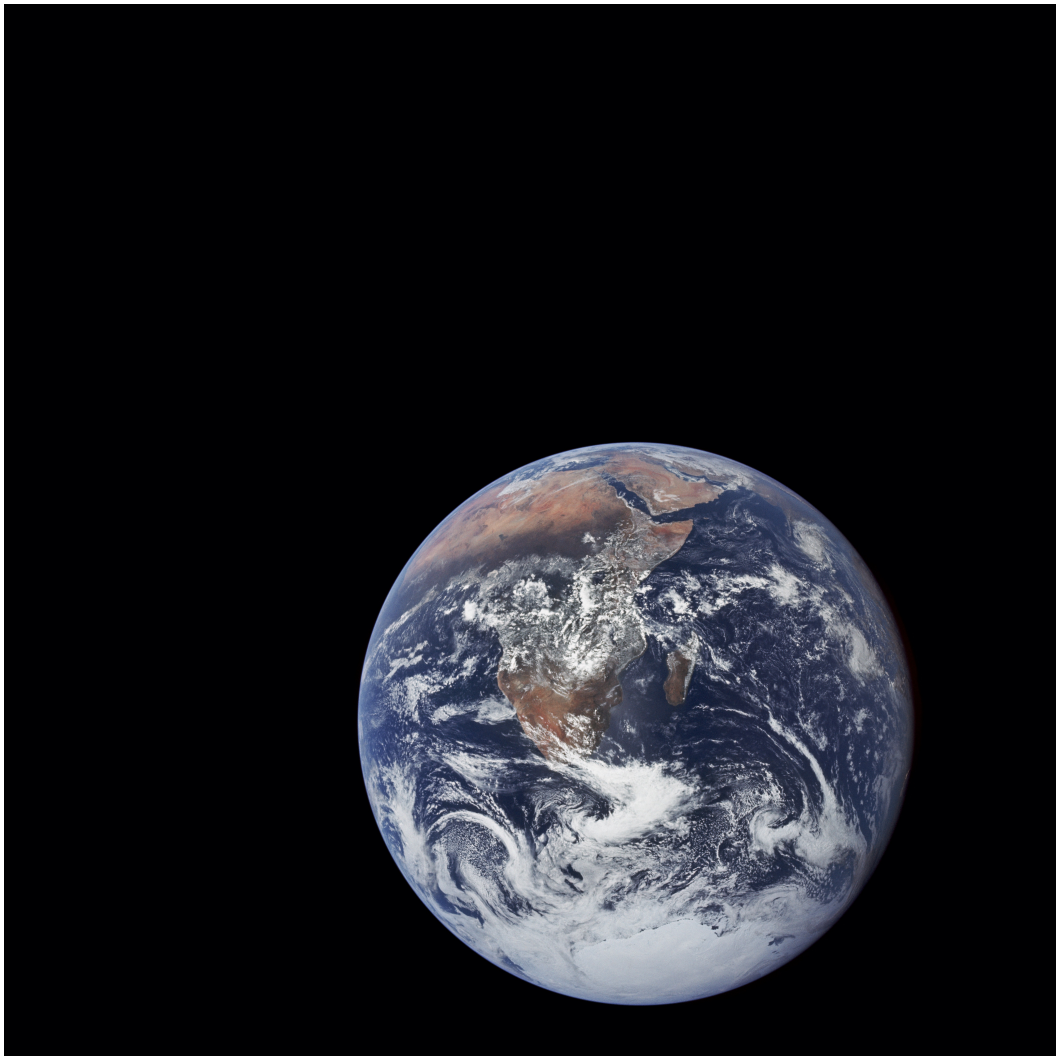


Fig. 7.27: The "Blue Marble" photograph of the Earth, taken by the crew of Apollo 17 five hours into their flight to the Moon on December 7, 1972, at about a distance of 29,000 km. NASA index AS17-148-22727.

Overcoming the challenges in human spaceflight during that era led to countless technological breakthroughs that literally changed the world in many ways, as well as fulfilling the rather banal geopolitical goals behind it. Initially an afterthought, the photographs of the Earth captured by astronauts on their journeys to the Moon and back gave a first-person view from that once unattainable position, and went as far as becoming symbols of the then nascent environmental movement — most famously the "Blue Marvel" picture taken on the last human flight to the Moon Figure 7.27 [Pet11]. The particular flight path of the Apollo 17 spacecraft allowed the astronauts to capture the almost completely illuminated surface of the Earth, as commander Eugene Cernan expressed at that moment: "And I suppose we're seeing as 100 percent full Earth as we'll ever see; certainly as I've ever seen. [...] it's about 99 percent pure. Bob, it's these kind of views that stick with you forever."

The *Blue Marble* [Wue12] and *Earthrise*, a similar view as shown in Figure 7.26 taken on the earlier Apollo 8 mission [Poo08] became celebrated and examined for their "totally unexpected impact," which "transformed thinking about the Earth and its environment," as "we saw ourselves and our place in the universe with new clarity." Seeing the whole Earth[3] inside the small field of view afforded by the windows in the capsules, astronauts got to experience what was later termed the "overview effect" by Frank White [Whi14]:

> The Overview Effect is a cognitive shift in awareness reported by some astronauts and cosmonauts during spaceflight, often while viewing the Earth from orbit, in transit between the Earth and the moon, or from the lunar surface. It refers to the experience of seeing firsthand the reality that the Earth is in space, a tiny, fragile ball of life, "hanging in the void," shielded and nourished by a paper-thin atmosphere. The experience often transforms astronauts' perspective on the planet and humanity's place in the universe. Some common aspects of it are a feeling of awe for the planet, a profound understanding of the interconnection of all life, and a renewed sense of responsibility for taking care of the environment.

This seems to confirm Socrates' speculation of only being able to recognizing the world for what it really is from a certain distance. While the ongoing research into this overview effect [Yad+16] was started as a series of interviews with astronauts and continues with efforts to reproduce it in VR [SQR19], the view from deep space is only the epitome of this effect. Similar reactions are recorded in astronauts visiting space stations or partaking in other missions with much lower orbits than the lunar transfers of the Apollo missions—from the altitude of the International Space Station, only about 3% of the Earth's surface can be seen at any time, as opposed to the about 40% seen in *Blue Marble*, or the 48% that could be seen from the distance of the Moon. White thus generalizes his description of the overview effect to refer to viewing common landscapes from far above, for which even tall mountaintops and airplanes can qualify, even if to an obviously much smaller effect.

The inverse stereographic projection as presented in this thesis, [4] attempts to merge the overview that can be obtained from a position high above with a focused and grounded view of details in an environment. NASA astronaut Sandra Magnus, who spent a total of

[3] The seminal counterculture magazine *Whole Earth Catalog*, foundational in modern environmentalism [Pep19] and "cyberculture" [Tur10] takes its name and the cover of its first edition from a satellite image mosaic taken in 1967 showing one of the earliest pictures recorded of the "whole Earth." Its founder Stewart Brand had to campaign NASA to release the then-rumored image, as the potential impact of such views has not yet been fully realized. The second and third covers of the magazine featured Earthrise.

[4] The technique itself builds on a principle first described by Hipparchus, another ancient Greek, living a quarter of a millennium after Socrates, who as the "greatest astronomer of antiquity" among other achievements also described and measured the orbit of the Moon and the Earth, for which he made the first known trigonometric tables.

157 days in space said the following to White regarding her experiences with the overview effect [Whi14]:

> It is kind of like a fractal system. If you are in a forest, you might be differentiating between the pine trees, the maple trees, and the oak trees, but if you see it from an airplane looking down at the forest, you just see trees. If you are in space, you just see the planet. It really is a fractal system, though I had not thought of it that way before. You take the resolution up a notch and your perception changes.

Focus+Context in general aims to combine these different steps in perception — from a great overview, down to fine details — into a single view, and the inverse stereographic projection technique achieves this through a simple geometric approach. A "perception change" was in fact observed in multiple people who tried the implemented system, with some reactions bordering on awe. As one stated: "this completely changes how I see the city."

Being so simple in its principle in using the most basic geometric shapes and operations, the inverse stereographic projection can be adapted to many applications across domains, and its implementation for virtual reality allows for intuitive interaction with the projected environments and the projection itself. The methods initially developed for adjusting and navigating the projection on the sphere could directly be ported over to the cone and cylinder for situations in which the sphere might not in fact be the "ideal shape." Even the bending technique, having nothing to do with the stereographic principle, was easily adapted into this implementation for use with the same interaction methods, simplifying its use more than the previous work that introduced it did. The result is a fully flexible focus+context system adaptable to various environments and use-cases, in which different projection or bending schemes can be switched seamlessly and interacted with in the same ways.

This flexibility also brings with it a vast array of aspects and parameters with it that demand a deeper and more focused investigation beyond the initial impressions and technical characteristics discussed here. Already at this point however, it stands to be a powerful tool for focus+context visualization that fully leverages most of the advantages that immersive environments offer, especially virtual reality: direct interaction with 3D controllers, better depth perception, and the large, empty workspace that can be filled with any arrangement. In contrast to desktop environments which act at best as a small window into the virtual world and whatever is displayed there, or previous focus+context methods which only the context in one direction even if implemented in VR, the inverse stereographic projection and its derivatives fully envelop the viewer, from having the area in focus at the feet, to projection the context upwards, possibly up to the zenith.



Fig. 7.28: Jackson Pollock in his studio in 1950 in a photograph taken by Hans Namuth.

In what could be seen as a historical parallel to the difference between desktop viewing of maps and the interaction with them in immersive environments, aided by the here presented system for focus+context, Jackson Pollock described in 1951 his at that time "radical" new painting method (Figure 7.28), which similarly stood in contrast to traditionally working a small easel in one's front [NFF51; Box17]:

> My painting is direct. I usually paint on the floor. I enjoy working on a large canvas. I feel more at home, more at ease, in the big area. Having the canvas on the floor, I feel nearer, more a part of the painting. This way, I can walk around it, work from all four sides, and be in the painting, similar to the Indian sand painters of the West.

In summary: the inverse stereographic projection technique was initially developed as a method for the stated purpose of this thesis — visualizing multilayered urban data — by adding one additional layer of geospatial information over the 3D city model in the otherwise unused space in the sky, and refining the way it is represented through deformations, ultimately leading to the presented projections on spheres, cones, and cylinders. Amalgamating long-established, as well as cutting-edge methods from cartographic projection, multi-perspective view deformation, and immersive analytics, the final implementation adheres to all design considerations set out in section 5.2.5:

- A 3D city model can be shown in all familiar perspectives, from a pedestrian to a top-down view. Acting as the focus area, it is not altered, except for a possible overlay of additional layers, such as e. g. transportation networks.
- The context hidden by occlusion or perspective distortion is revealed in a legible way by projecting it to a sphere or other shape, without creating discontinuities. The unavoidable distortions can be balanced as necessary in trade-offs with the visibility of distant objects, or the preservation of the appearance of straight lines.
- The focus+context view can be adjusted using simple real-time controls that are universally applicable to all deformation schemes.
- Immersive technologies are fully utilized, not only for the interaction and navigation methods that are possible with spatially tracked controllers, but also for the fully enveloping view they provide, which can be completely and usefully filled out with the projections, resulting in a panoramic view in all directions, hopefully heightening the sense of immersion in more than just the technical way.

Having accomplished this, its main achievements however could be seen as the new methods and improvements it presents for geospatial focus+context, which may at the same time be simpler to use and more effective in their results than comparable techniques. Its potential applications exceed the scope of this thesis by appearing to be a very promising method for geospatial exploration, and especially navigation, whether of a virtual environment, or the real world through an overlay of augmented reality.

Part IV

Conclusions and Future Work

8 Conclusions

Socrates' reasoning about seeing "through the water" [Pla+23] could also be broadly seen as a general metaphor on visualization itself. Like "lifting their heads out of the water" to escape murkiness, visualization is fundamentally intended to provide clarity for seeing certain information that may be hidden otherwise, and to provide overviews, literally through different perspectives, or figuratively through agglomeration of information.

In investigating the uses immersive environments could have for visualizing multilayered geospatial urban data, two distinct methods were developed that fully take advantage of these environments to allow new ways of seeing the connections between different layers of data, and the connections between distant and close areas. In other words: a different, possibly more coherent overview over multiple layers of geospatial information, and a new method for an overview over focus+context. Both approaches could find uses beyond visualizing city maps and models.

Seen on a very high level, agglomerating geospatial data can on go into breadth and depth, as is the case with most other kind of information. Here, the first obviously means covering larger areas, while the second can mean different things: finer details/higher resolution data of the same area, different/additional kinds of data, or data from different points in time. The last two dimensions can be seen and organized as layers of a map, and they usually require special considerations for how to manage their visualization with sometimes complex arrangements, or complex cartographic design to merge them into single views.

With immersive environments, the theoretically available visual space is vastly larger than a regular monitor or even multiple printouts. This large space, together with the added benefit of stereoscopic vision can be taken advantage of to handle spatial data layers in innovative and hopefully more intuitive ways that would otherwise be unpractical. Reducing the constraint of space allows layers to take up more of it without interfering with one another, and invites to find uses for visual space that has previously been underutilized, like the sky.

The MCV stack (section 6) uses that space by separating and vertically spacing out multiple map layers, in a way that aims to preserve a sense of connection between them, while also evoking the most common illustration of the concept of multiple layers. The inverse stereographic projection and its related methods (section 7) can also be seen as

using separation to show connectedness: information outside its focus area is gradually being separated from its original position on the terrain, and lifted up into the sky to become visible again.

The implementations of these two ideas were meant to investigate the hypotheses made in the second part of the thesis. An evaluation of the MCV stack in comparison with other MCV methods (blitting and a grid of layers) using a task derived from research in urbanism attempted to answer the hypothesis from section 4.4, repeated again below:

> The dynamic layer stack could enhance multilayer map comparisons, compared to alternative methods such as regular juxtaposition of these layers in a grid, or temporal juxtaposition / blitting.

The results did not show a clear-cut confirmation or rejection, but rather a "depends." Analysis revealed how the MCV stack *can* indeed be better in multiple aspects for *some* users, but not all, and pointed to the kinds of user characteristics that correlate with a preference for or a better performance with any particular system. Offering users a choice of presentation method, and allowing the methods to be more flexible to better fit individual users appeared to be more important than attempting to optimize any particular system. What could also be observed however was a general enthusiasm for VR as a medium to present and interact with the multiple layers, irrespective of MCV method, and coming both from novice users as well as from urbanists. This plays into the first, more general hypothesis made in section 3.2:

> The virtual space and interactivity afforded by immersive environments offers solutions to visualizing multilayered urban data at multiple scales, while being highly engaging.

The implementation and evaluation of the MCV Stack (and its comparable methods) have already demonstrated most of this to be true, with only "multiple scales" not being directly addressed through them, other than by more or less traditional methods of zooming. This part is what the inverse stereographic projection is more directly aimed at, being a novel method for visualizing geospatial focus+context.

Regarding engagement, the projection of the city map on the sphere received more direct enthusiasm from those who tried it in during informal testing. Where the MCV systems mostly benefited from being shown inside a still quite novel environment for most users, who were delighted to use VR even for mundane tasks, they are comparatively regular map views, just arranged differently in virtual space. The enveloping, inverse stereographic projection on the other hand provoked utterances such as "it completely changes the way I see the city" even from experienced urbanists, and its unusual nature invited users to experiment with different control schemes to explore the cityscape in

various, often playful, ways. It appears to have exceeded the "engagement boost," or the "wow-effect" provided from something just being shown in VR, and shows genuinely interesting and novel views of urban environments, which almost invariably evoke a comparison to certain scenes in Christopher Nolan's 2010 film *Inception*, even when viewed on a regular desktop screen or printout.

In being so engaging, and utilizing virtual reality for the full extent of the surrounding view it provides and the interactions that are possible with the 3D-tracked controllers, at least the last part of the hypothesis made in section 5.4 appears to have be confirmed by its implementation and informal evaluation:

> The inverse stereographic projection with its variations is a powerful tool for exploration of potentially multilayered focus+context in diverse geospatial applications, which can fully exploit the benefits of immersive environments while being intuitive to understand.

The remainder of this hypothesis has however only been argued in theory or demonstrated informally. During its development, the system meant to host the inverse stereographic projection to a sphere evolved to include not only other projection surfaces, but also "competing" focus+context methods that rely on geometry bending instead of projection, using the same interaction schemes originally created just for adjusting the parameters of the sphere. In doing so, it became a flexible and highly adjustable system, allowing for the host of potential applications touched upon in section 7.4. Even a single possible use case, walked through in section 7.4.1 could employ multiple projection methods and of course constantly changing parameters within the span of minutes, if not seconds.

Evaluating only the projection on the sphere, e. g. in comparison to other surfaces or the bending methods, or even restricting it to certain settings for its parameters would severely limit the flexibility and therefore usefulness of the integrated system, distorting its perception. As Besançon et al. wrote in *The State of the Art of Spatial Interfaces for 3D Visualization* [Bes+21]: "it is essential to consider the unique affordances of interaction paradigms in different scenarios or contexts" — it would rather make more sense to test the same, fully-functional system under many various use-cases and scenarios, and analyze emergent user behavior. They highlighted this evaluation of complex visualization systems as being a major challenge, and observed this about the techniques they surveyed:

> Indeed, these techniques are designed mostly for experts and researchers in specific domains. Domain experts using data visualization for sense-making rarely evaluate visualizations based on completion time or error-based measures but rather on the potential to turn data into insights. [...]

> Since 3D visualization techniques often target a specific audience of domain experts, the alternative evaluation strategies that are most common focus on feedback from these users. Recruiting experts can be difficult, but the feedback they provide in user studies is invaluable. [Bes+21]

Seen this way, the above hypothesis regarding the inverse stereographic projection has already been mostly validated even with just the feedback gathered from experts in the informal evaluations using the scenario of urban exploration. What remains to further strengthen the case are more rigorous studies involving other use-cases, other domains, and potential enhancements to the system, especially the only briefly mentioned multi-layer capability of the focus+context system. These will however have to be relegated to future work, as the overall question of this thesis, posed in section 2.4:

> How can immersive technologies be employed to improve the visualization of multilayered urban data, and make multiple scales easier to navigate?

appears to already have been satisfactorily answered at this point, with the demonstration and evaluation of the systems in section 6: Virtually Physical Data Layer Stack and section 7: Virtual Data Sphere for Inverse Stereographic Projection as two examples. Immersive environments have been shown to be more than capable of supporting such visualizations and interactions with them that would be neither feasible nor as compelling with traditional methods.

9 Future Work

To limit the scope of this thesis for brevity and focus, and because of time constraints, many avenues for further research, implementation, evaluation, comparison, and discussion had to be relegated to future work. Investigation of e. g. cartographic styles and generalization techniques, navigation and interaction, and the application of the implemented techniques beyond urban environments had to be strictly limited in order to stay within these bounds, as well as further explorations of possible variations and combinations of these methods, which can appear endless.

9.1 Further Evaluation

9.1.1 Formal Evaluation of the Inverse Stereographic Projection

The inverse stereographic projection has so far only been informally evaluated, in a flexible and fully interactive system that also incorporated its related focus+context methods as being only one toggle away. Being a novel projection scheme, its characteristics do warrant more rigorous investigations by themselves, where the parameters are locked, and different settings or techniques could be compared against each other. Basic orientation, navigation, and spatial memory tasks, combinations of these, and other methods commonly employed in the evaluation of 3D and geospatial visualization techniques [Bes+21] could be used here. Different projection surfaces or bending methods could be compared against each other similar to how the MCV systems were in section 6.2, and more importantly be used across different use-cases to discover possible advantages each variation may have for certain scenarios.

The technically conformal deformation of the inverse stereographic projection to a sphere should be evaluated on a perceptual basis: how distorted does it appear to viewers? Which shapes of which sizes suffer the most from it? What configurations of the projection mitigate or exacerbate these effects? How does it compare against the non-conformal projections on a cone or cylinder, or against the geometry-bending techniques?

The perception of scale is another interesting avenue for investigation. Currently, zooming inside the system works by simply scaling the 3D representation of the city, and

keeping it anchored to the floor of the virtual room, which results it appearing as anything between an architectural scale model and a first-person streetview simulation. Piumsomboon et al. [Piu+18] have been investigating the influence of altering the interpupillary distance of the virtual cameras for urban visualization and telepresence, which can result in changes to the perception of user's size and therefore their relation to the city model: do they see themselves as a giant walking around toy-sized buildings, or as "Superman" flying above a real skyline? What implications could such perceptions have for the projections of the context area?

9.1.2 Continued Evaluation of the MCV Stack

Likewise, the MCV Stack and its related methods could benefit from further investigation and evaluation. Even the data recorded during the first study (section 6.2) still contains a wealth of information that could be further analyzed, such as fine-grained temporal information about where participants were looking at and what they were doing. Figure 9.1 shows the data for participant №24's gaze, where each color represents a sample of time during which any particular map layer was being looked at during the test sessions.

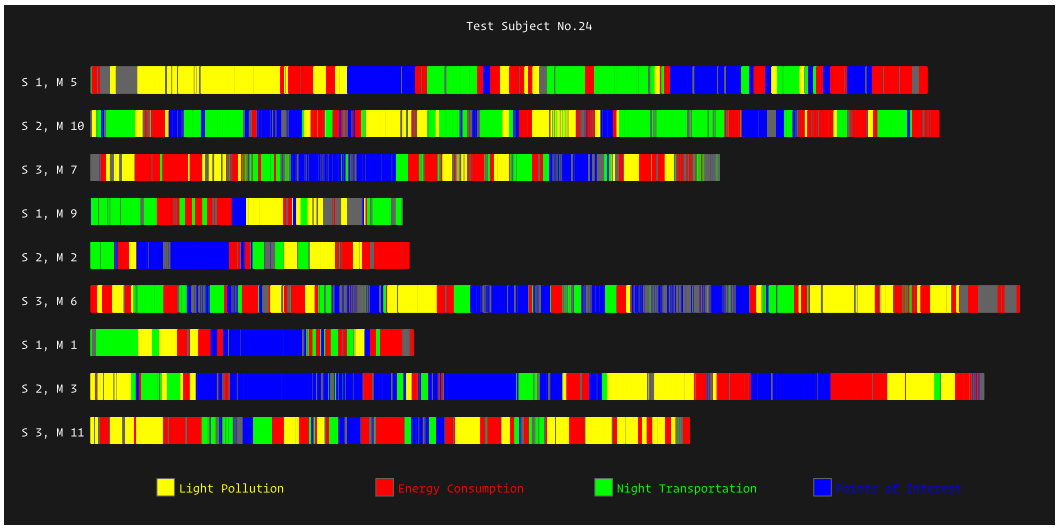


Fig. 9.1: An excerpt of the gaze data recorded during the study in section 6.2, showing which map layers participant №24's looked at in fine detail for each scenario.

The stack method in the configuration shown in section 6.2 has been shown to work for its intended purpose and compared favorably against a grid or a blitting method for certain users. Different configurations warrant closer inspections, similarly to the inverse stereographic projection and its various parameters. At which point does the legibility of tilted layers decrease beyond usefulness? How much does this depend on the resolution of HMDs, or on the kind of data being shown?

The study (section 6.2) also only compared the three MCV methods in the most "fair" configuration of four layers, in which each system is reasonably usable. At which point

will one system more clearly prevail when more layers are added, if any? Does this depend on scenarios and use-cases? How do different spatial arrangements within one of these MCV systems influence their usability?

Being situated in a virtual environment, all three MCV methods could be enhanced to show more than just layers of maps floating in space. They could include full 3D models of buildings and structures on their maps, and floating and billboarding labels and icons extending from them. Being tilted, the stack method could also show advantages over the other methods in visualizing such three-dimensional scenes. Conversely, the other methods could also be tilted from their vertical position they had in first study for 3D visualizations, and comparisons could be made as to which system is better suited for such, and in which configurations.

9.2 Enhancements

As already hinted at above, there are certain enhancements that can be done to the presented systems, all of which could warrant further investigations into optimal configurations and subsequent (comparative) evaluations. Some possibilities and ideas are listed below.

9.2.1 Inverse Stereographic Projection Enhancements

Several possible enhancements to the focus+context projection methods are immediately imaginable. The “flattening” before projection (section 5.3.5) of building heights and terrain to mitigate extreme distortions or exaggerations could be automatically handled in way that neither is perceptible in the projection. This could be achieved through an automatic regulation of the flattening parameter dependent on distance and projection, or by extruding fully flattened objects back towards the center of the sphere (or perpendicularly from any surface) to a scaled height, calculated from the amount of compression they underwent. This could weaken the observed structure-from-motion effect, but would result in a more natural appearance. In fact, the second method is already being employed here for the implementation of the geometry-bending approach, thus almost ready to be ported to the part of the shader responsible for the stereographic projection.

Another method from the geometry-bending technique that could be adapted to the projection would be a smooth transition zone to replace the sharp intersection between the original terrain and the projection surface. Although perhaps looking smoother, such a transition could overall be detrimental in other ways to the perception of the deformed context, as previously discussed (section 7.3.3).

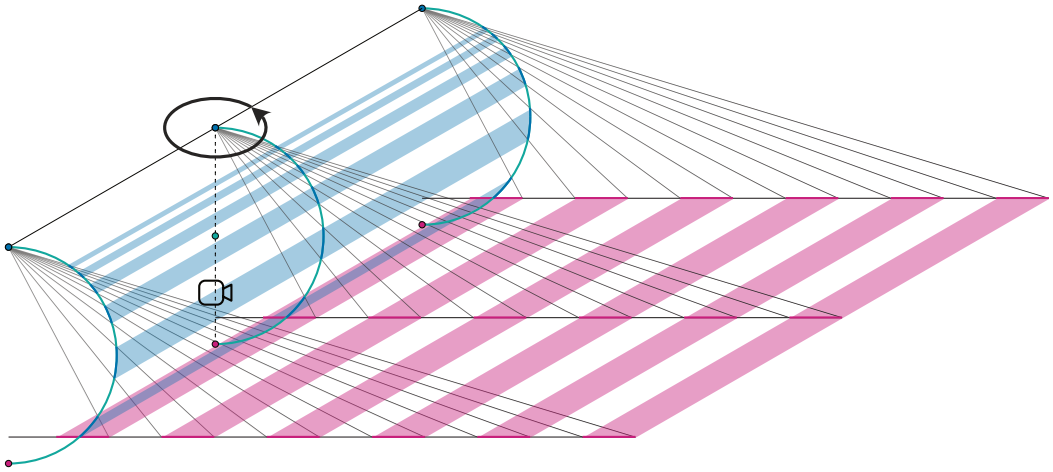


Fig. 9.2: An application of the inverse stereographic projection to a vertically-oriented half-cylinder, with the projection origin not being a single point above the viewer, but rather the infinite line at the top of the cylinder. Lines perpendicular to the viewing direction remain straight in their projection.

Finally the preservation of straight lines from geometry-bending that takes place along a straight line as demonstrated by Lorenz et al. [Lor+08], which in the implementation by Pasewaldt et al. [Pas+14] follows the camera's direction could also be achieved for another variant of the stereographic projection, at least for lines that are perpendicular to that direction. Figure 9.2 shows a projection to a vertically-oriented half-cylinder, which rotates with the gaze of the viewer — (some) lines remain straight, while the compression of distant context takes place in the same gradual way as in the projection to a sphere.

9.2.2 MCV Stack Enhancements

The different kinds of spatial arrangements that might be necessary to accommodate more than four layers could be seen as a first "enhancement" for the MCV stack system. Figure 10.12 in the annex shows three variants of stacking for six layers; either all in one tall column, or in two separate stacks, which may or may not be directly facing the viewer, depending on whether being seen at the exact same angles is more important than saving some space.

Side-by-side stacks could also be used to show the same data layers from different points in time: an arc could be formed with a number of available snapshots for quick spatiotemporal comparisons, or a full "carousel" for cyclically, e. g. weekly repeating data. A button press should bring any stack with desired point in time to the front, with the remaining columns following accordingly. This kind of interaction and automatic arrangement on an arc is already implemented in the MCV stack system, along with a scheme for adjusting the vertical order of layers within a stack: selecting any two layers, and initiating their switch.

Also already noted above, the MCV methods could further benefit from the possibilities immersive environments offer by displaying fully three-dimensional city models or other objects over the map layers. The MCV metaphor could also be extended further by providing additional coordinated views of smaller focus areas, such as pedestrian views of individual streets and their buildings in separate views. These views could take the shape of additional "floating windows," or in a *world-in-miniature* visualization [SCP95; ESB21], where the layers provide an overview over a larger area that can be explored from a first-person perspective.

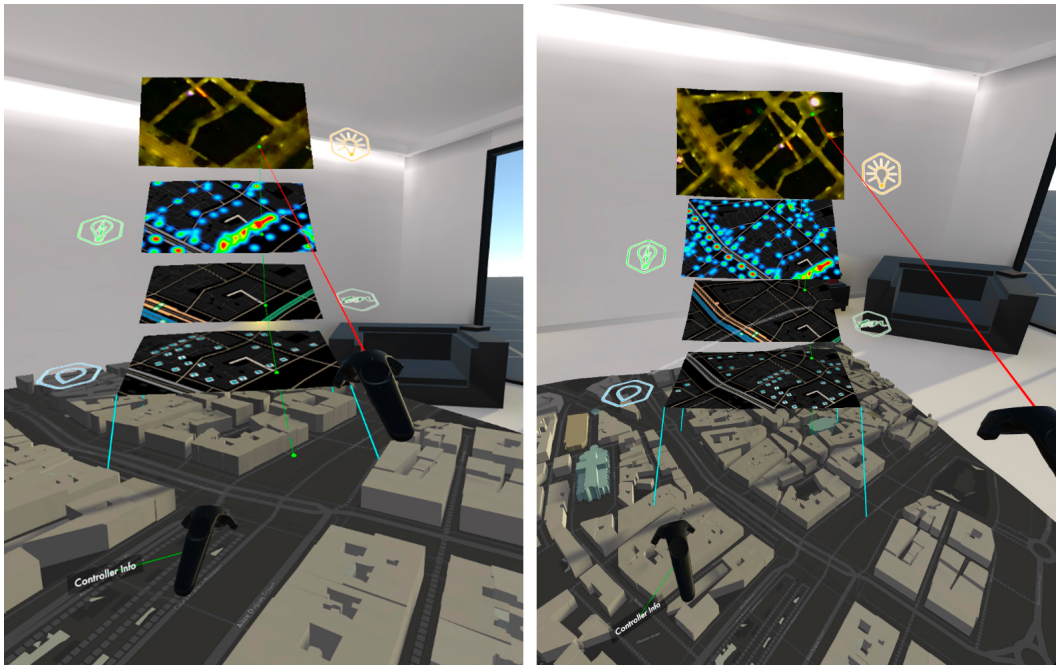


Fig. 9.3: The MCV stack adapted to act as a overview+detail visualization over a 3D city model. Shown on the left as a model floating at table-height, and on the right as lying on the (virtual) floor. Blue lines from the corner of the lowest layer show from where in the model the layers are projected from.

Such a world-in-miniature view, or the simpler "floating windows" into certain areas on the maps would be akin to an inverted overview+detail visualization, where the "detail" is the virtual environment surrounding the MCV system. A simpler possibility overview+detail can be achieved by inverting this arrangement: a larger model of the city, perhaps the whole region is laid on the floor, with the stack of layers floating above a certain area of focus. A first attempt at such a system is shown in Figure 9.3.

9.2.3 Hybrid: Overview+Detail + Focus+Context

This second kind of overview+detail, where the stack floats over a portion of a large city model, could also serve as a starting point for a hybrid between the MCV stack system and the focus+context system based on the inverse stereographic projection. The stacked layers would then act as a "hyperfocus" area, and an the immediate region

around it, similar to what is seen displayed in 3D in Figure 9.3 would be the "regular" focus area, or view extent, in terms of the stereographic projection. Behind that, the context would curve up to the sky as is possible with the different projection schemes and their parameters. An additional feature possible in such a hybrid would be the stack acting as a selection of available layers for projection on the sphere. So far, only the primary and secondary projection parameters are implemented for real-time interaction (section 7.1.2) — the stack could show and enable a selection of pre-defined sets of tertiary projection and visualization parameters for visualizing the context.

Of course, the same merging of the two systems would also work with the world-in-miniature version described above, which could rather be seen as an addition of the miniature stack to a the focus+context projection. A possible implementation could be for augmented reality, where the sphere or cone is used in an urban setting to show immediate context and further-away connections, and the stack acts as both a handheld, multi-layered mini-map which shows either the same focus area or a different extent, and as a selector for how to project context, as described above.

This hybridization between the MCV stack and the projection of context, and in fact between overview+detail and focus+context goes back to the genesis and motivation for terrain deformation (section 5.1), which lead to the idea for the inverse stereographic projection, where a view from "inside" the stack was desired. Such a combination of both systems could lead to a more holistic answer to the question from section 2.4 than either could on their own:

> How can immersive technologies be employed to improve the visualization of multilayered urban data, and make multiple scales easier to navigate?

Part V

Annex and Bibliography

Annex 10

10.1 Gestalt Laws

The following figures are a selection of the so-far discovered Gestalt Laws as discussed in section 2.1.4, mostly adapted from [War13].

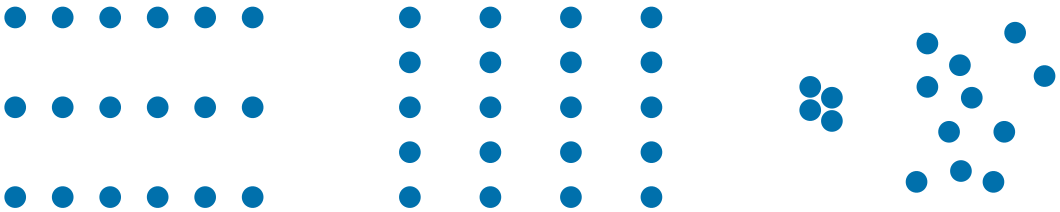


Fig. 10.1: Proximity: The matrices of dots are perceived as groups of rows or columns, just as the clumps of dots is perceived as consisting of tho groups because of spatial proximity or concentration.

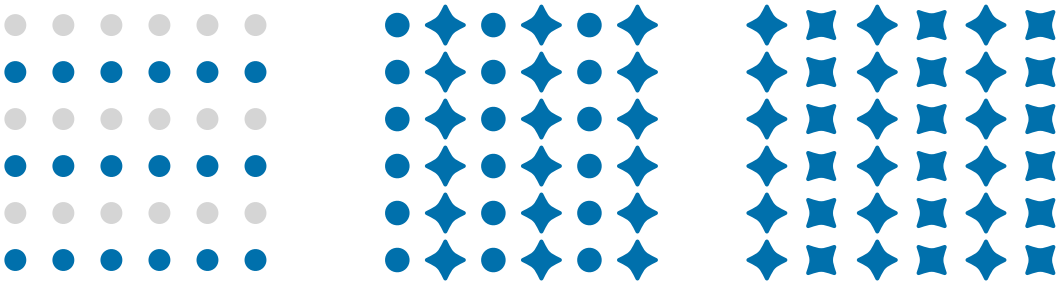


Fig. 10.2: Similarity: The matrix elements are perceived as rows or columns because of similar colors, shapes or orientations.

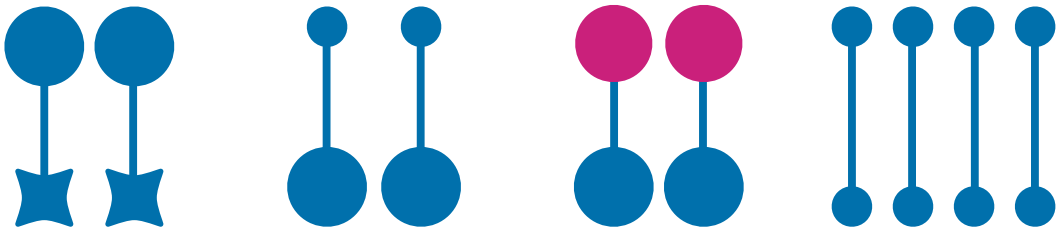


Fig. 10.3: Connectedness: Elements that are close together or similar are not perceived as a group as much as those that are connected.

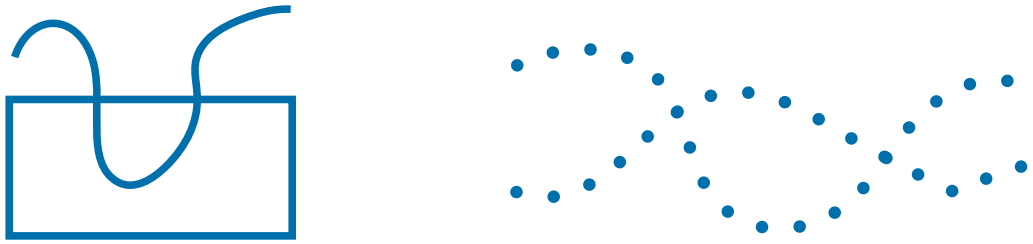


Fig. 10.4: Continuity: Intersecting lines are perceived to continue their smooth shape and are not assumed to be taking sharp corners or be divided.



Fig. 10.5: Symmetry: The same identical shapes appear more closely grouped when being on an axis of symmetry.



Fig. 10.6: Closure: The red shape appears as a circle occluded by the rectangle and not as a broken ring.



Fig. 10.7: Relative Size: The smaller blue areas are more likely to be perceived as one object.

10.2 MCV Study Introduction

The following page contains the document all participants in the study who comparatively evaluated the immersive MCV systems (section 6.2) received at the very beginning of their session as a help them familiarize themselves with the scenario and the tasks. A translation into French was also provided upon request. These pages were kept available during the whole experiment as a reference, in addition to the tooltips provided inside the virtual environment.

Introduction

Dear participant,

Welcome to our study about visualizing and understanding urban data concerning **public lighting**!

In this study, we're interested in **your opinions**, and also about how we can best present you information necessary to make decisions.

Please read this introduction carefully to make sure you understand the test procedure.

SCENARIO

The City of Nantes wants *you* to help reduce **energy consumption** and **light pollution** from publicly-owned streetlights.

Reducing consumption is good for the climate and the budget, and reducing light pollution is important for nocturnal animals (including fish), for helping people sleep, and for preserving a natural dark sky at night.

Of course, there are tradeoffs, and it's obviously not desirable to just turn all lights off. Active nightlife areas benefit from more light, as do busy streets and transportation hubs.

DIRECT PUBLIC WORK

We have highlighted a number of areas in different regions in Nantes where work could be done to dim the lights or install newer, more efficient lamps, but we need *you* to decide which ones you think would benefit the **most** from a change.

To make an informed choice on where to *focus public work*, you will need to understand different **layers** of data — basically, different maps of the same places showing different information about them.

INFORMATION LAYERS

Each region has the following **layers** of information available:

1. **Light Pollution**:
   A photograph taken from a plane, that shows how brightly the city is illuminated at night.
2. **Energy Consumption:**
   A map showing the energy consumption of each public lamp on the streets.
3. **Night Transport:**
   A map that shows all public transport lines that operates at night, as well as *Bicloo* stations.
4. **Night Life:**
   A map of restaurants, bars, hotels, sights and other places that people may visit at night.

These maps show the information during a typical Friday night between April and August — most offices are closed, and people meet in the city, or go home later; many using public transport.

Introduction

TASK: CHOICE

You will be presented with different **regions** of the city, with three highlighted smaller areas in each. You explore all layers of these regions and then select **one** of the three **candidate areas**, based on how you understand and weigh the information you see. Which area of the three do you think is **most** in need of change? There are no wrong choices — we need your honest opinion!

After you make your selection, we would just like you to write down a very short explanation on **why** you chose it. It could be just a few words or a short sentence, like “there is a too high energy consumption here,” or “there is no need for such amount of light considering there is no public transportation here”.

VISUALIZATION SYSTEMS

We are also testing three different systems to show these layers to you, to see which one is most helpful for understanding all at once. All systems work in Virtual Reality (VR) — you will put on a VR headset and interact with two VR controllers.

One controller allows you to **navigate** the maps — zoom, pan, point at it, or switching layers. The other controller lets you highlight and **select** from the three given areas to compare. You can choose which hand you use for which, but we recommend to use your dominant hand for the map navigation tools.

For each region where we want you to make your choice of one of three areas, you will be presented the layers in one of three different systems. At the beginning of the test, you will get an **in-depth tutorial** on general interaction with one system, and then two short ones for the remaining systems.

After the tutorials, you will just use alternating systems with new regions — **each system 3 times!**

PROCEDURE

After a quick visual acuity and color vision test, you will find the VR devices and a PC to fill out questionnaires in the test room. Before the first choice tasks, the questionnaire asks you a few general questions, and then instructs you to proceed to the VR system.

Each choice task starts with a calibration of the system: just follow the white point with your eyes until you see the virtual room and the maps. After you explored the layers and made your choice, you will be instructed to take off the VR headset and fill out a questionnaire about the task you just did. You then put on the headset again and repeat this until you've finished all the tasks.

After the last task you fill out the questionnaire a last time, some general questions about ALL tasks and systems will follow — then you're done!

The test should take less than 1:30 hours — the longest part is the tutorial! The rest is very quick.

If you have any question at any time during the experiment or while filling out forms, please don't hesitate to ask the experimenter. We want you to feel as confident and comfortable as possible!

10.3 Saccade Polar Plots

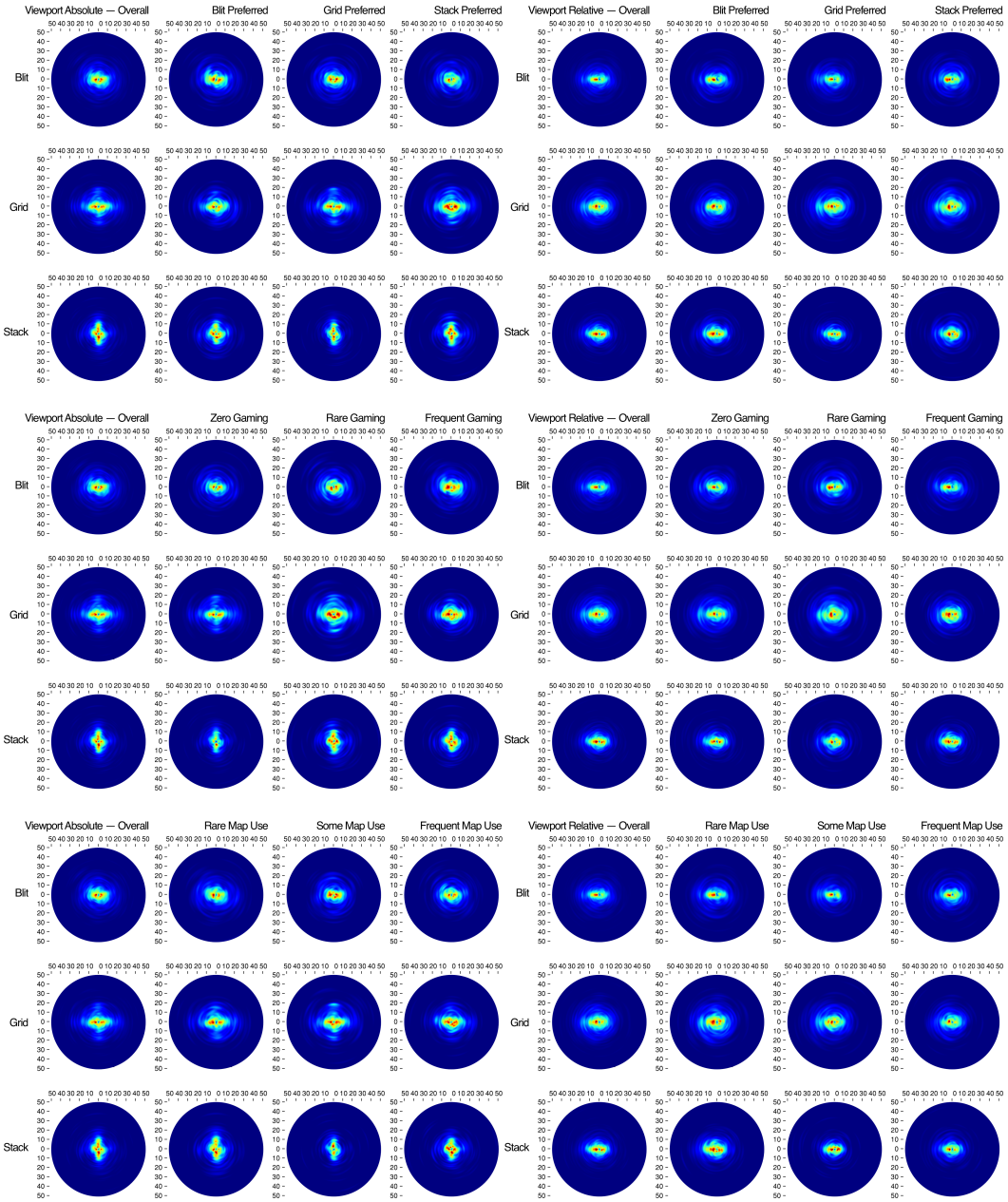


Fig. 10.8: Polar plots showing the density probability distributions of absolute and relative saccades in the viewport—the eye movement component—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**).

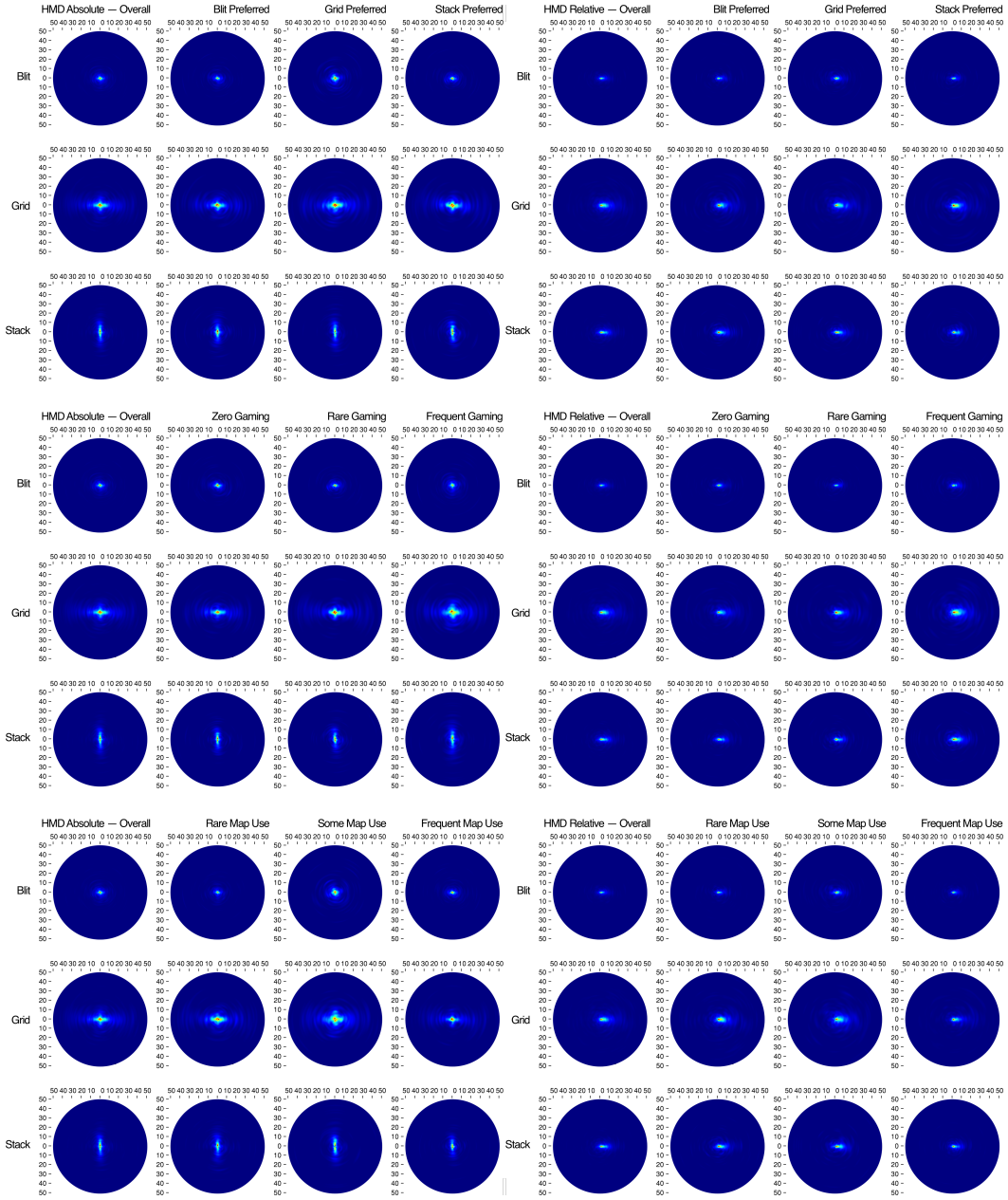


Fig. 10.9: Polar plots showing the density probability distributions of absolute and relative saccades of the head—the HMD movement component—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**).

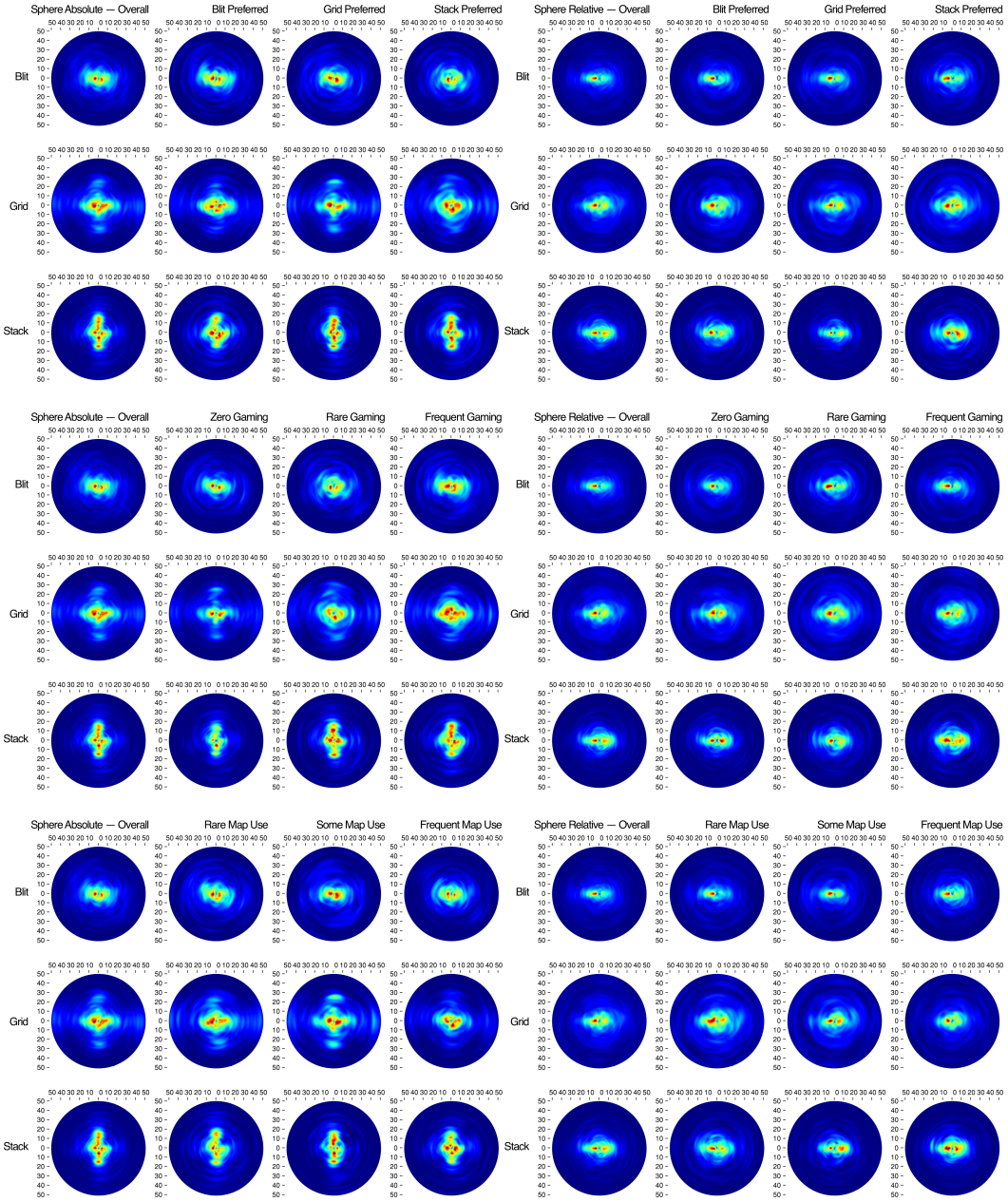


Fig. 10.10: Polar plots showing the density probability distributions of absolute and relative total saccades—the gaze, as projected on a sphere—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**).

10.4 Additional MCV Configurations

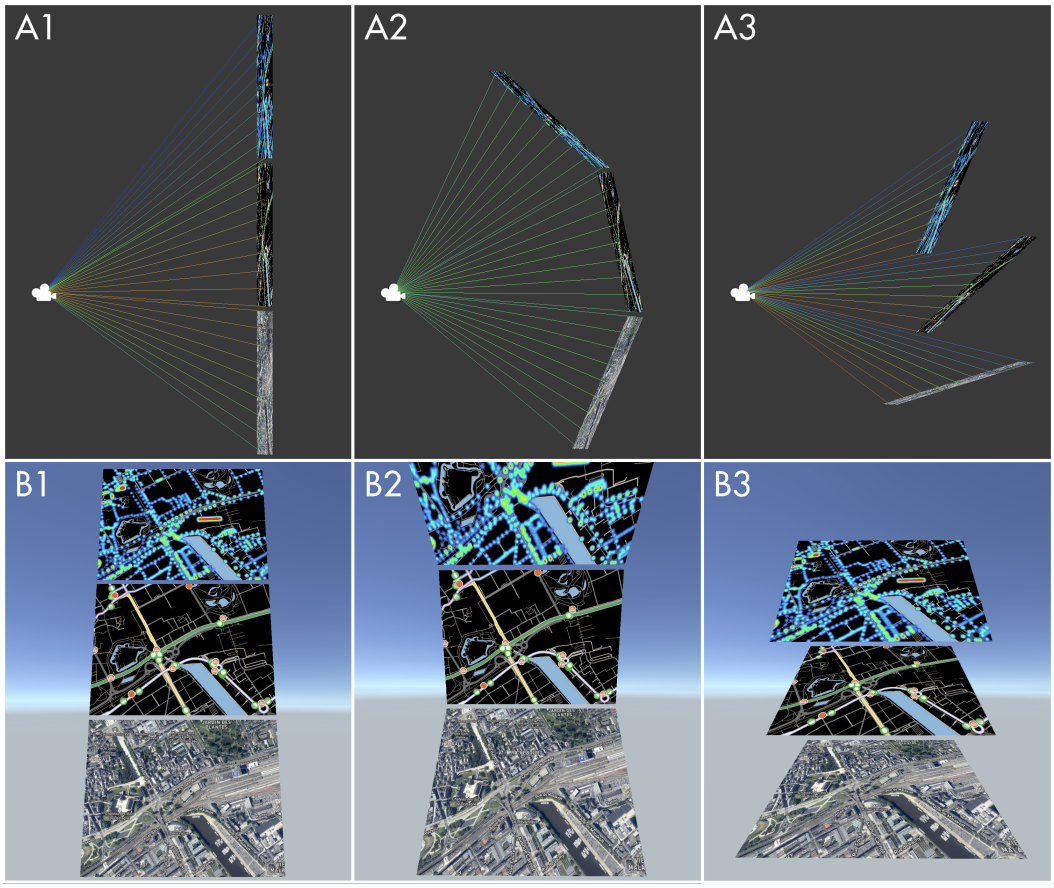


Fig. 10.11: (**A**bove: Isometric views from the side of a flat configuration of three layers, a stack with each layer perpendicular to the viewline (2), and a stack with each layer tilted at 40 degrees (3). Colored lines between illustrate differences in distance between camera and equidistant points on the map views: green is the average distance, a shift to red signifies less distance, a shift to blue more. (**B**)elow: the corresponding camera views from (**A**), with equal fields of view (tilted up in (**B1**)). All layers equal in size.

Figure Figure 10.11 shows how the stack system compares against a flat grid of map views regarding distortion and their use of vertical space. Three views of identical size are used to illustrate the visual differences when the views are (a) arranged vertically and coplanar to each other; (b) equidistant from the viewer, and each perpendicular (90 degrees) to the line extending from the viewer to the center of each view (viewline); and (c) arranged in a stack similar to (b), but the views are titled by 40 degrees to the viewline.

As seen by the distances between viewer and points on the map views in Figure 10.11, perspective distortion can be unevenly distributed across views in a coplanar arrangement (1), as each has a different distance and tilt toward the viewer. Using the stack system in an extreme setting to keep each layer perpendicular and equidistant (2), perspective distortions are minimal and each layer appears in the same shape to the viewer in VR—the distortions in the camera views are due to projection to a flat screen image.

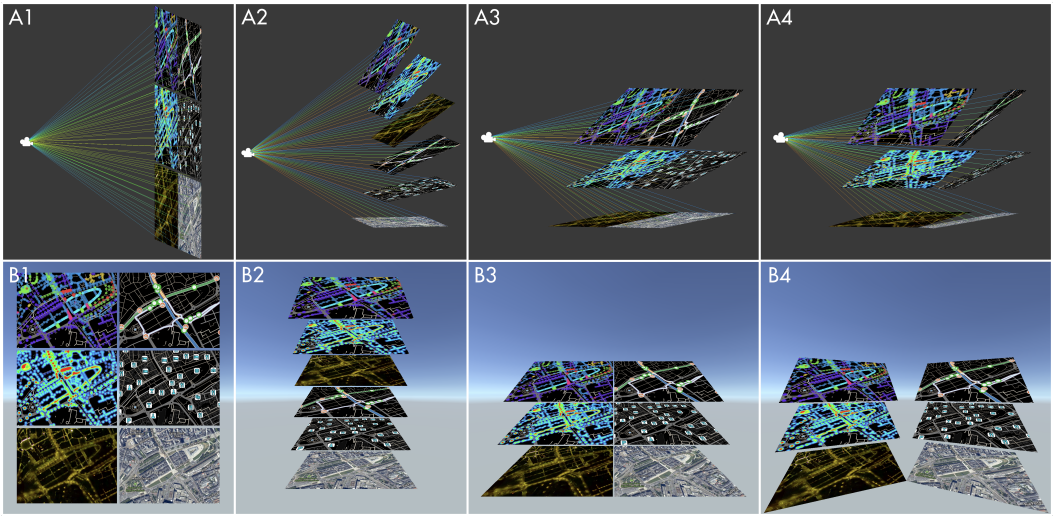


Fig. 10.12: (**A**)bove: Isometric views from the side of a flat configuration of six layers (1), one stack with each layer tilted at 25 degrees (2), two stacks with each layer tilted at 35 degrees (3), and two stacks as in (3) but rotated to face the viewer (4). Colored lines illustrate distances as in 10.11. (**B**)elow: the corresponding camera views from (**A**), with equal fields of view (cropped sides in (**B1** and (**B2**)). All layers equal in size.

This does however take up more of the field of view; it can be seen by the top layer being cut off in B2. A more moderate setting of the stack system, with the layers being tilted from perpendicularity (3) reduces the visual space all layers take up, while adding more perspective distortion uniformly to all layers.

Figure 10.12 shows comparisons between a grid and different stack arrangements for six layers, similarly to Figure Figure 10.12.A 2-by-3 grid (1) takes up similar vertical space as a the same layers all arranged vertically in a stack of tilted layers (2), but more horizontal space. Multiple stacks can be arranged side-by-side to form a hybrid between grid and stack—similar to the "shelves" proposed in [Liu+20], but with the stack system keeping each layer tilted to the same angle. (4) shows the full functionality of the two stacks side-by-side, keeping a constant distance and view angle for all layers, and a reduction in functionality that constrains the layers to all face in the same direction. (3) takes up less horizontal space than (4), at the cost of distorting each layer differently, while (4) keeps the stack system's full functionality of equal distortion of each layer at the cost of requiring more horizontal space and more distance between the stacks to avoid overlap.

10.5 Shadergraph

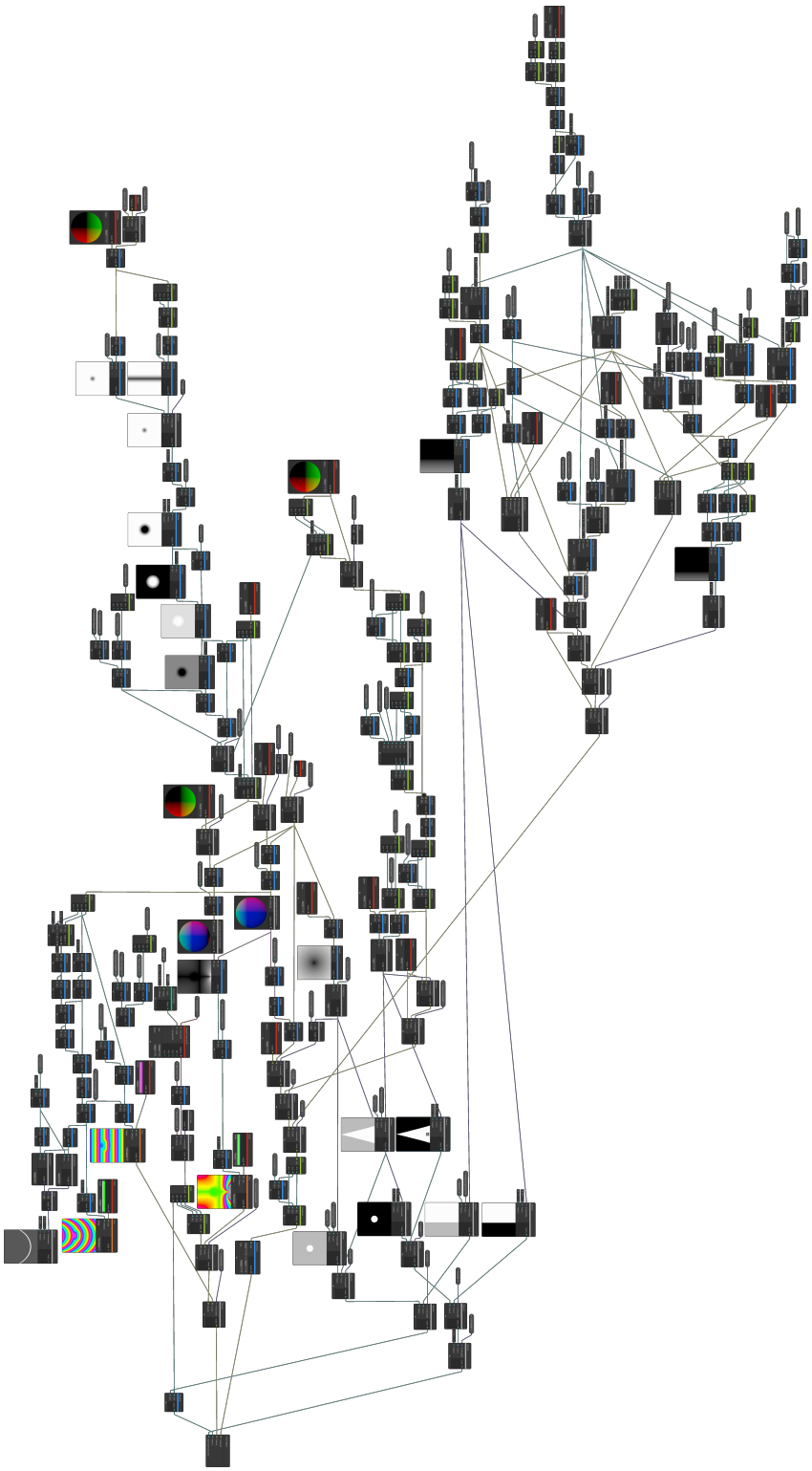


Fig. 10.13: The shader graph implemented for the inverse stereographic projection on a sphere, cone, and cylinder, with all projection parameters listed in section 7.1.2 and an implementation of the geometry bending approach from [Lor+08].

Bibliography

[AA10] Natalia Adrienko and Gennady Adrienko. “Spatial generalization and aggregation of massive movement data”. In: *IEEE Transactions on visualization and computer graphics* 17.2 (2010), pp. 205–219 (cit. on p. 36).

[ABP18] A Afrooz, H Ballal, and C Pettit. “IMPLEMENTING AUGMENTED REALITY SANDBOX IN GEODESIGN: A FUTURE FOR GEODESIGN.” In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 4.4 (2018) (cit. on p. 37).

[AEN10] Christopher Andrews, Alex Endert, and Chris North. “Space to Think: Large High-Resolution Displays for Sensemaking”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 55–64 (cit. on p. 48).

[Alo+18] Luis Alonso, Yan Ryan Zhang, Arnaud Grignard, et al. “Cityscope: a data-driven interactive simulation tool for urban design. Use case volpe”. In: *International conference on complex systems*. Springer. 2018, pp. 253–261 (cit. on pp. 32, 37).

[Ama+10] MP Amado, CV Santos, EB Moura, and VG Silva. “Public participation in sustainable urban planning”. In: *International journal of human and social sciences* 5.2 (2010), pp. 102–108 (cit. on p. 6).

[AN12] C. Andrews and C. North. “Analyst’s Workspace: An embodied sensemaking environment for large, high-resolution displays”. In: *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 2012, pp. 123–131 (cit. on p. 56).

[And+07] Gennady Andrienko, Natalia Andrienko, Piotr Jankowski, et al. “Geovisual analytics for spatial decision support: Setting the research agenda”. In: *International journal of geographical information science* 21.8 (2007), pp. 839–857 (cit. on p. 36).

[And+11] Christopher Andrews, Alex Endert, Beth Yost, and Chris North. “Information visualization on large, high-resolution displays: Issues, challenges, and opportunities”. In: *Information Visualization* 10.4 (2011), pp. 341–355. eprint: https://doi.org/10.1177/1473871611415997 (cit. on p. 48).

[AQ20] Samuel Ang and John Quarles. “GingerVR: An Open Source Repository of Cybersickness Reduction Techniques for Unity”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE. 2020, pp. 460–463 (cit. on p. 21).

[Att+11] Simon Attfield, Gabriella Kazai, Mounia Lalmas, and Benjamin Piwowarski. “Towards a science of user engagement (position paper)”. In: *WSDM workshop on user modelling for Web applications*. 2011, pp. 9–12 (cit. on p. 25).

[Bar10] Matthias Bartelmann. “Gravitational lensing”. In: *Classical and Quantum Gravity* 27.23 (2010), p. 233001 (cit. on p. 68).

[Bat+14] Sarah E. Battersby, Michael P. Finn, E. Lynn Usery, and Kristina H. Yamamoto. “Implications of Web Mercator and Its Use in Online Mapping”. In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 49.2 (June 2014), pp. 85–101 (cit. on p. 165).

[Bat13] Michael Batty. *The new science of cities*. MIT press, 2013 (cit. on p. 31).

[Bat17] Michael Batty. “The future journal”. In: *Environment and Planning B: Urban Analytics and City Science* 44.1 (Jan. 2017), pp. 6–9 (cit. on p. 28).

[Bat18] Michael Batty. “Artificial intelligence and smart cities”. In: *Environment and Planning B: Urban Analytics and City Science* 45.1 (2018), pp. 3–6. eprint: https://doi.org/10.1177/2399808317751169 (cit. on p. 32).

[Bat19] Michael Batty. “Urban analytics defined”. In: *Environment and Planning B: Urban Analytics and City Science* 46.3 (Mar. 2019), pp. 403–405 (cit. on p. 28).

[BC04] Emily Brown and Paul Cairns. “A grounded investigation of game immersion”. In: *CHI'04 extended abstracts on Human factors in computing systems*. 2004, pp. 1297–1300 (cit. on p. 24).

[Ber+15] Joshua Bergerson, Ralph T. Muehleisen, W. Bo Rodda, et al. “Designing Future Cities: LakeSIM Integrated Design Tool For Assessing Short And Long Term Impacts Of Urban Scale Conceptual Designs”. In: *ISOCARP Review* 11 (Jan. 2015) (cit. on p. 162).

[Ber67] Jacques Bertin. “Sémiologie graphique: Les diagrammes-Les réseaux-Les cartes”. In: (1967) (cit. on pp. 25, 35).

[Bes+13] Anne-Kathrin Bestgen, Dennis Edler, Frank Dickmann, and Lars Kuchinke. “Grid or no grid: distance distortion in recognizing spatial information from complex cartographic maps”. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 35. 35. 2013 (cit. on p. 55).

[Bes+21] Lonni Besançon, Anders Ynnerman, Daniel F. Keefe, Lingyun Yu, and Tobias Isenberg. “The State of the Art of Spatial Interfaces for 3D Visualization”. In: *Computer Graphics Forum* 40.1 (Feb. 2021), pp. 293–326 (cit. on pp. 179–181).

[BH14] Michael Batty and Andrew Hudson-Smith. “Visual analytics for urban design”. In: *Urban Design* 132 (2014), pp. 38–41 (cit. on p. 29).

[BHB99] Woodrow Barfield, Claudia Hendrix, and Karl-Erik Bystrom. “Effects of stereopsis and head tracking on performance using desktop virtual environment displays”. In: *Presence: Teleoperators & Virtual Environments* 8.2 (1999), pp. 237–240 (cit. on p. 24).

[BHS09] Gordon Bell, Tony Hey, and Alex Szalay. “Beyond the data deluge”. In: *Science* 323.5919 (2009), pp. 1297–1298 (cit. on p. 34).

[BHY18] Alexander Bock, Charles Hansen, and Anders Ynnerman. “OpenSpace: Bringing NASA Missions to the Public”. In: *IEEE Computer Graphics and Applications* 38.5 (2018), pp. 112–118 (cit. on p. 166).

[Bil15] Ibrahim Oztug Bildirici. “Quasi indicatrix approach for distortion visualization and analysis for map projections”. In: *International Journal of Geographical Information Science* 29.12 (2015), pp. 2295–2309. eprint: https://doi.org/10.1080/13658816.2015.1074236 (cit. on p. 137).

[BJW17] Jonas Buddeberg, Bernhard Jenny, and Wesley Willett. “Interactive shearing for terrain visualization: an expert study”. In: *GeoInformatica* 21.3 (July 2017), pp. 643–665 (cit. on p. 86).

[BL94] Michael Batty and Paul A Longley. *Fractal cities: a geometry of form and function*. Academic press, 1994 (cit. on p. 31).

[Bla+18] Karl Bladin, Emil Axelsson, Erik Broberg, et al. “Globe Browsing: Contextualized Spatio-Temporal Planetary Surface Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (Jan. 2018), pp. 802–811 (cit. on p. 165).

[BM07] Doug A Bowman and Ryan P McMahan. “Virtual reality: how much immersion is enough?” In: *Computer* 40.7 (2007), pp. 36–43 (cit. on p. 24).

[BM12] Richard Brath and Peter Macmurchy. “Sphere-based information visualization: Challenges and benefits”. In: *2012 16th International Conference on Information Visualisation*. IEEE. 2012, pp. 1–6 (cit. on p. 79).

[BM13] Matthew Brehmer and Tamara Munzner. “A multi-level typology of abstract visualization tasks”. In: *IEEE transactions on visualization and computer graphics* 19.12 (2013), pp. 2376–2385 (cit. on p. 27).

[BN05] Robert Ball and Chris North. “Effects of Tiled High-Resolution Display on Basic Visualization and Navigation Tasks”. In: *CHI ’05 Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’05. Portland, OR, USA: Association for Computing Machinery, 2005, pp. 1196–1199 (cit. on p. 47).

[BNB07] Robert Ball, Chris North, and Doug A. Bowman. “Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’07. San Jose, California, USA: Association for Computing Machinery, 2007, pp. 191–200 (cit. on p. 48).

[Bon+09] Leonardo Bonanni, Xiao Xiao, Matthew Hockenberry, et al. “Wetpaint: Scraping through Multi-Layered Images”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’09. Boston, MA, USA: Association for Computing Machinery, 2009, pp. 571–574 (cit. on p. 58).

[Böt+08] Joachim Böttger, Martin Preiser, Michael Balzer, and Oliver Deussen. “Detail-In-Context Visualization for Satellite Imagery”. In: *Computer Graphics Forum* 27.2 (2008), pp. 587–596. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2008.01156.x (cit. on pp. 69, 70, 73, 75, 79, 91, 157).

[Box17] Sarah Boxer. “The Photos That Changed Pollock’s Life”. In: *The New York Times* (Dec. 15, 2017) (cit. on p. 172).

[Bra+16] Mickaël Brasebin, Sidonie Christophe, Florence Jacquinod, Anouk Vinesse, and Hortense Mahon. “3D Geovisualization & stylization to manage comprehensive and participative Local Urban Plans”. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2016), p. 83 (cit. on pp. 22, 65, 75).

[Bra84] Ralph A. Bradley. “14 Paired comparisons: Some basic procedures and examples”. In: *Nonparametric Methods*. Vol. 4. Handbook of Statistics. Elsevier, 1984, pp. 299–326 (cit. on p. 111).

[BTW17] Monica Billger, Liane Thuvander, and Beata Stahre Wästberg. “In search of visualization challenges: The development and implementation of visualization tools for supporting dialogue in urban planning processes”. In: *Environment and Planning B: Urban Analytics and City Science* 44.6 (2017), pp. 1012–1035 (cit. on p. 6).

[Büs+18] Wolfgang Büschel, Jian Chen, Raimund Dachselt, et al. “Interaction for immersive analytics”. In: *Immersive Analytics*. Springer, 2018, pp. 95–138 (cit. on pp. 27, 49).

[Buy+17] Ismail Buyuksalih, Serdar Bayburt, Gurcan Buyuksalih, et al. “3D MODELLING AND VISUALIZATION BASED ON THE UNITY GAME ENGINE–ADVANTAGES AND CHALLENGES.” In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 4 (2017) (cit. on p. 95).

[BW05] Stephen Brooks and Jacqueline L. Whalley. “A 2D/3D Hybrid Geographical Information System”. In: *Proceedings of the 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*. GRAPHITE ’05. Dunedin, New Zealand: Association for Computing Machinery, 2005, pp. 323–330 (cit. on p. 53).

[BW08] Stephen Brooks and Jacqueline L. Whalley. “Multilayer hybrid visualizations to support 3D GIS”. In: *Computers, Environment and Urban Systems* 32.4 (2008). Geographical Information Science Research – United Kingdom, pp. 278–292 (cit. on pp. 53, 61).

[BWK20] Pauline Bimberg, Tim Weissker, and Alexander Kulik. “On the Usage of the Simulator Sickness Questionnaire for Virtual Reality Research”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2020, pp. 464–467 (cit. on p. 122).

[BYB18] Lauren E Buck, Mary K Young, and Bobby Bodenheimer. “A comparison of distance estimation in HMD-based virtual environments with different HMD-based conditions”. In: *ACM Transactions on Applied Perception (TAP)* 15.3 (2018), pp. 1–15 (cit. on p. 21).

[Can19] Galo Canizares. “Another Allegory of Simulation”. In: *Journal of Architectural Education* 73.1 (2019), pp. 103–106. eprint: https://doi.org/10.1080/10464883.2019.1560815 (cit. on p. 80).

[Car03] M Sheelagh T Carpendale. “Considering visual variables as a basis for information visualisation”. In: (2003) (cit. on p. 35).

[CB17] Carlos Carbonell Carrera and Luis A Bermejo Asensio. “Landscape interpretation with augmented reality and maps to improve spatial orientation skill”. In: *Journal of Geography in Higher Education* 41.1 (2017), pp. 119–133 (cit. on p. 37).

[CC18] Oakyoon Cha and Sang Chul Chong. “Perceived Average Orientation Reflects Effective Gist of the Surface”. In: *Psychological science* 29.3 (2018), pp. 319–327 (cit. on p. 22).

[CCN14] Paul Cairns, Anna Cox, and A Imran Nordin. “Immersion in digital games: review of gaming experience research”. In: *Handbook of digital games* 1 (2014), p. 767 (cit. on p. 24).

[CDT02] Dempsey Chang, Laurence Dooley, and Juhani E Tuovinen. “Gestalt theory in visual screen design: a new look at an old subject”. In: *Proceedings of the Seventh world conference on computers in education conference on Computers in education: Australian topics-Volume 8*. Australian Computer Society, Inc. 2002, pp. 5–12 (cit. on p. 13).

[ÇFL10] Arzu Çöltekin, Sara Irina Fabrikant, and Martin Lacayo. “Exploring the efficiency of users’ visual analytics strategies based on sequence analysis of eye movement recordings”. In: *International journal of geographical information science* 24.10 (2010), pp. 1559–1575 (cit. on p. 36).

[Cha+15] Tom Chandler, Maxime Cordeil, Tobias Czauderna, et al. “Immersive Analytics”. In: *2015 Big Data Visual Analytics (BDVA)*. IEEE Computer Society. 2015, pp. 1–8 (cit. on p. 26).

[Cha11] Samuel Challéat. “La mise en débats des territoires de la lumière”. In: *Journée d’études sur Les effets de la participation*. GIS Démocratie & Participation, École des Hautes Études en Sciences Sociales (EHESS). Paris, France, Oct. 2011 (cit. on p. 102).

[Che+14] M. Chen, S. Walton, K. Berger, et al. “Visual Multiplexing”. In: *Computer Graphics Forum* 33.3 (2014), pp. 241–250. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.12380 (cit. on p. 56).

[Che+17] Zhutian Chen, Yifang Wang, Tianchen Sun, et al. “Exploring the design space of immersive urban analytics”. In: *Visual Informatics* 1.2 (2017), pp. 132–142 (cit. on p. 75).

[CKB09] Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. “A Review of Overview+detail, Zooming, and Focus+context Interfaces”. In: 41.1 (Jan. 2009) (cit. on pp. 67–69).

[CL18] Min Chen and Hui Lin. “Virtual geographic environments (VGEs): originating from or beyond virtual reality (VR)?” In: *International Journal of Digital Earth* 11.4 (2018), pp. 329–333. eprint: https://www.tandfonline.com/doi/pdf/10.1080/17538947.2017.1419452 (cit. on pp. 38, 39).

[Cla82] James H Clark. “The geometry engine: A VLSI geometry system for graphics”. In: *acm siggraph Computer Graphics* 16.3 (1982), pp. 127–133 (cit. on p. 25).

[ÇLZ16] A Çöltekin, I Lokka, and M Zahner. “On the usability and usefulness of 3D (geo) visualizations–A focus on virtual reality environments”. In: *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 41 (2016) (cit. on p. 37).

[CM00] Andy Cockburn and Bruce McKenzie. “An evaluation of cone trees”. In: *People and Computers XIV—Usability or Else!* Springer, 2000, pp. 425–436 (cit. on p. 26).

[CM01] Andy Cockburn and Bruce McKenzie. “3D or not 3D? Evaluating the effect of the third dimension in a document management system”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2001, pp. 434–441 (cit. on p. 26).

[CM84] William S Cleveland and Robert McGill. “Graphical perception: Theory, experimentation, and application to the development of graphical methods”. In: *Journal of the American statistical association* 79.387 (1984), pp. 531–554 (cit. on p. 25).

[CMK18] Tom Chandler, Thomas Morgan, and Torsten Wolfgang Kuhlen. “Exploring immersive analytics for built environments”. In: *Immersive Analytics*. Springer, 2018, pp. 331–357 (cit. on p. 27).

[CMS88] Michael Chen, S. Joy Mountford, and Abigail Sellen. “A Study in Interactive 3-D Rotation Using 2-D Control Devices”. In: *Proceedings of the 15th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’88. New York, NY, USA: Association for Computing Machinery, 1988, pp. 121–129 (cit. on p. 79).

[Çöl+20] Arzu Çöltekin, Amy L Griffin, Aidan Slingsby, et al. “Geospatial Information Visualization and Extended Reality Displays”. In: *Manual of Digital Earth*. Springer, Singapore, 2020, pp. 229–277 (cit. on pp. 31, 33, 38–40).

[Cor+19] Maxime Cordeil, Andrew Cunningham, Benjamin Bach, et al. “IATK: An Immersive Analytics Toolkit”. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 200–209 (cit. on p. 95).

[Cos+21] Jonathas Costa, Alexander Bock, Carter Emmart, et al. “Interactive Visualization of Atmospheric Effects for Celestial Bodies”. In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2021), pp. 785–795 (cit. on p. 166).

[CP19] Jeremy Clifton and Stephen Palmisano. “Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality”. In: *Virtual Reality* (2019), pp. 1–16 (cit. on p. 148).

[Cri+15] Verina Cristie, Matthias Berger, Peter Bus, Ashwani Kumar, and Bernhard Klein. “CityHeat: Visualizing Cellular Automata-Based Traffic Heat in Unity3D”. In: *SIGGRAPH Asia 2015 Visualization in High Performance Computing*. SA ’15. Kobe, Japan: Association for Computing Machinery, 2015 (cit. on p. 95).

[CRM91] Stuart K Card, George G Robertson, and Jock D Mackinlay. “The information visualizer, an information workspace”. In: *Proceedings of the SIGCHI Conference on Human factors in computing systems*. 1991, pp. 181–186 (cit. on pp. 25, 26).

[Cru+92] Carolina Cruz-Neira, Daniel J Sandin, Thomas A DeFanti, Robert V Kenyon, and John C Hart. “The CAVE: audio visual experience automatic virtual environment”. In: *Communications of the ACM* 35.6 (1992), pp. 64–73 (cit. on p. 26).

[CSD93] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. “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 ’93. Anaheim, CA: Association for Computing Machinery, 1993, pp. 135–142 (cit. on p. 49).

[CT05] Kristin A Cook and James J Thomas. *Illuminating the path: The research and development agenda for visual analytics*. Tech. rep. Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2005 (cit. on pp. 23, 25, 26).

[Cui+10] J. Cui, P. Rosen, V. Popescu, and C. Hoffmann. “A Curved Ray Camera for Handling Occlusions through Continuous Multiperspective Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), pp. 1235–1242 (cit. on p. 68).

[CV95] James E. Cutting and Peter M. Vishton. “Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth*”. In: *Perception of Space and Motion*. Ed. by William Epstein and Sheena Rogers. Handbook of Perception and Cognition. San Diego: Academic Press, 1995, pp. 69–117 (cit. on p. 19).

[DA34] Charles Henry Deetz and Oscar Sherman Adams. *Elements of map projection*. 4th ed. US Government Printing Office, 1934 (cit. on p. 81).

[Dav+18] Erwan J. David, Jesús Gutiérrez, Antoine Coutrot, Matthieu Perreira Da Silva, and Patrick Le Callet. “A Dataset of Head and Eye Movements for 360° Videos”. In: *Proceedings of the 9th ACM Multimedia Systems Conference*. MMSys ’18. Amsterdam, Netherlands: Association for Computing Machinery, 2018, pp. 432–437 (cit. on p. 111).

[Dav+19] Erwan Joël David, Pierre Lebranchu, Matthieu Perreira Da Silva, and Patrick Le Callet. "Predicting artificial visual field losses: A gaze-based inference study". In: *Journal of Vision* 19.14 (Dec. 2019), pp. 22–22. eprint: https://arvojournals.org/arvo/content_public/journal/jov/938293/i1534-7362-19-14-22.pdf (cit. on pp. 111, 119).

[Day94] Alan Day. "From map to model: the development of an urban information system". In: *Design Studies* 15.3 (1994), pp. 366–384 (cit. on p. 27).

[DeL+10] Bertrand DeLongueville, Alessandro Annoni, Sven Schade, Nicole Ostlaender, and Ceri Whitmore. "Digital earth's nervous system for crisis events: real-time sensor web enablement of volunteered geographic information". In: *International Journal of Digital Earth* 3.3 (2010), pp. 242–259 (cit. on p. 32).

[Dev+18] Alexandre Devaux, Charlotte Hoarau, Mathieu Brédif, and Sidonie Christophe. "3D Urban Geovisualization: In Situ Augmented And Mixed Reality Experiments." In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 4.4 (2018) (cit. on p. 37).

[DH09] S Oliver Daum and Heiko Hecht. "Distance estimation in vista space". In: *Attention, Perception, & Psychophysics* 71.5 (2009), pp. 1127–1137 (cit. on pp. 19, 20).

[DiB+92] David DiBiase, Alan M MacEachren, John B Krygier, and Catherine Reeves. "Animation and the role of map design in scientific visualization". In: *Cartography and geographic information systems* 19.4 (1992), pp. 201–214 (cit. on p. 35).

[DMK05] Jason Dykes, Alan M. MacEachren, and Menno Jan Kraak. "Introduction Exploring Geovisualization". English (US). In: *Exploring Geovisualization*. Elsevier Ltd, Dec. 2005, pp. 1–19 (cit. on p. 34).

[DMP16] Laurent Dupont, Laure Morel, and Marc Pallot. "Exploring the Appropriateness of Different Immersive Environments in the Context of an Innovation Process for Smart-Cities". In: *2016 International Conference on Engineering, Technology and Innovation/IEEE International Technology Management Conference (ICE/ITMC)*. IEEE. 2016, pp. 1–9 (cit. on p. 40).

[Dor+18] Harish Doraiswamy, Juliana Freire, Marcos Lage, Fabio Miranda, and Claudio Silva. "Spatio-Temporal Urban Data Analysis: A Visual Analytics Perspective". In: *IEEE Computer Graphics and Applications* 38.5 (Sept. 2018), pp. 26–35 (cit. on pp. 22, 29, 40).

[Dou04] Paul Dourish. *Where the Action is: The Foundations of Embodied Interaction (Bradford Books)*. New edition. MIT Press, 2004 (cit. on p. 56).

[Dro97] Peter Droege. *Intelligent environments: spatial aspects of the information revolution*. Elsevier, 1997 (cit. on p. 31).

[Dun+17] David Dunn, Cary Tippets, Kent Torell, et al. "Wide field of view varifocal near-eye display using see-through deformable membrane mirrors". In: *IEEE transactions on visualization and computer graphics* 23.4 (2017), pp. 1322–1331 (cit. on p. 21).

[Dwy+18] Tim Dwyer, Kim Marriott, Tobias Isenberg, et al. "Immersive analytics: An introduction". In: *Immersive analytics*. Springer, 2018, pp. 1–23 (cit. on pp. 23, 25, 27, 50).

[ED15] Dennis Edler and Frank Dickmann. "Elevating streets in urban topographic maps improves the speed of map-reading". In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 50.4 (2015), pp. 217–223 (cit. on p. 54).

[Edl+14] Dennis Edler, Oliver Huber, Claudia Knust, Manfred F Buchroithner, and Frank Dickmann. “Spreading Map Information over Different Depth Layers–An Improvement for Map-Reading Efficiency?” In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 49.3 (2014), pp. 153–163 (cit. on p. 55).

[EFI14] Barrett M Ens, Rory Finnegan, and Pourang P Irani. “The personal cockpit: a spatial interface for effective task switching on head-worn displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 3171–3180 (cit. on p. 49).

[EHI14] Barrett Ens, Juan David Hincapié-Ramos, and Pourang Irani. “Ethereal planes: a design framework for 2D information space in 3D mixed reality environments”. In: *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 2014, pp. 2–12 (cit. on p. 49).

[EI17] Barrett Ens and Pourang Irani. “Spatial analytic interfaces: Spatial user interfaces for in situ visual analytics”. In: *IEEE computer graphics and applications* 37.2 (2017), pp. 66–79 (cit. on p. 49).

[Ell87] James Elliot. *The City in Maps: urban mapping to 1900*. British Library, 1987 (cit. on p. 27).

[EM05] Laura Ermi and Frans Mäyrä. “Fundamental components of the gameplay experience: Analysing immersion”. In: *Worlds in play: International perspectives on digital games research* 37.2 (2005), pp. 37–53 (cit. on p. 24).

[Ens+21] Barrett Ens, Benjamin Bach, Maxime Cordeil, et al. “Grand Challenges in Immersive Analytics”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Vol. 17. New York, NY, USA: ACM, May 2021, pp. 1–17 (cit. on p. 100).

[Epi20] Epic Games. *Unity 3D*. Version 2019.2.9f1. Apr. 20, 2020 (cit. on p. 95).

[ESB21] David Englmeier, Wanja Sajko, and Andreas Butz. “Spherical World in Miniature: Exploring the Tiny Planets Metaphor for Discrete Locomotion in Virtual Reality”. In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Mar. 2021, pp. 345–352 (cit. on p. 185).

[Fer+13a] Nivan Ferreira, James T Klosowski, Carlos E Scheidegger, and Cláudio T Silva. “Vector field k-means: Clustering trajectories by fitting multiple vector fields”. In: *Computer Graphics Forum*. Vol. 32. 3pt2. Wiley Online Library. 2013, pp. 201–210 (cit. on p. 30).

[Fer+13b] Nivan Ferreira, Jorge Poco, Huy T Vo, Juliana Freire, and Cláudio T Silva. “Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips”. In: *IEEE transactions on visualization and computer graphics* 19.12 (2013), pp. 2149–2158 (cit. on pp. 29, 30).

[Fer+15] Nivan Ferreira, Marcos Lage, Harish Doraiswamy, et al. “Urbane: A 3D framework to support data driven decision making in urban development”. In: *2015 IEEE conference on visual analytics science and technology (VAST)*. IEEE. 2015, pp. 97–104 (cit. on pp. 29, 30).

[FH98] J. M. Flach and J. G. Holden. “The Reality of Experience: Gibson’s Way”. In: *Presence* 7.1 (1998), pp. 90–95 (cit. on p. 24).

[Fis18] Dafna Fisher-Gewirtzman. “Perception of density by pedestrians on urban paths: An experiment in virtual reality”. In: *Journal of Urban Design* 23.5 (2018), pp. 674–692 (cit. on p. 22).

[For89] Jay W Forrester. *Urban Dynamics*. MIT Press, Cambridge, MA, 1989, p. 285 (cit. on p. 31).

[FS69] R Buckminster Fuller and J Snyder. *Operating manual for spaceship Earth*. Southern Illinois University Press, 1969 (cit. on p. 31).

[FSN20] J. A. W. Filho, W. Stuerzlinger, and L. Nedel. "Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration". In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (2020), pp. 514–524 (cit. on pp. 50, 61).

[Fuh+05] Sven Fuhrmann, Paula Ahonen-Rainio, Robert M Edsall, et al. "Making useful and useable geovisualization: design and evaluation issues". In: *Exploring geovisualization*. Elsevier, 2005, pp. 551–566 (cit. on p. 34).

[GAM15] Tiago Gonçalves, Ana Paula Afonso, and Bruno Martins. "Cartographic visualization of human trajectory data: Overview and analysis". In: *Journal of Location Based Services* 9.2 (2015), pp. 138–166 (cit. on p. 22).

[Gam17] Phillipe Gambert. "Éclairage public à Nantes : des idées pour payer moins". In: *Ouest France* (2017) (cit. on p. 101).

[Gar74] Wendell R Garner. "The stimulus in information processing". In: *Sensation and measurement*. Springer, 1974, pp. 77–90 (cit. on p. 16).

[GC52] James J Gibson and Janet Cornsweet. "The perceived slant of visual surfaces—optical and geographical." In: *Journal of Experimental Psychology* 44.1 (1952), p. 11 (cit. on p. 20).

[Ger+07] Daniel M Germán, Lloyd Burchill, Alexandre Duret-Lutz, et al. "Flattening the Viewable Sphere." In: *Computational Aesthetics*. Citeseer. 2007, pp. 23–28 (cit. on p. 81).

[GF12] Amy L Griffin and Sara I Fabrikant. "More maps, more users, more devices means more cartographic challenges". In: *Cartographic Journal* 49.4 (2012), pp. 298–301 (cit. on p. 35).

[GGA20] Huadong Guo, Michael F. Goodchild, and Alessandro Annoni, eds. *Manual of Digital Earth*. Springer Nature, 2020 (cit. on pp. 31, 33, 39, 165).

[Gib50] James J Gibson. "The perception of the visual world." In: (1950) (cit. on p. 19).

[Gib79] James Jerome Gibson. *The ecological approach to visual perception*. Houghton, Mifflin and Company, 1979 (cit. on pp. 9, 23).

[GKS92] David V Gibson, George Kozmetsky, and Raymond W Smilor. *The technopolis phenomenon: Smart cities, fast systems, global networks*. Rowman & Littlefield, 1992 (cit. on p. 31).

[Gle+11] Michael Gleicher, Danielle Albers, Rick Walker, et al. "Visual comparison for information visualization". In: *Information Visualization* 10.4 (2011), pp. 289–309. eprint: https://doi.org/10.1177/1473871611416549 (cit. on pp. 59, 60).

[Gle18] M. Gleicher. "Considerations for Visualizing Comparison". In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (2018), pp. 413–423 (cit. on pp. 60–62, 103).

[GM92] Melvyn A Goodale and A David Milner. "Separate visual pathways for perception and action". In: *Trends in neurosciences* 15.1 (1992), pp. 20–25 (cit. on p. 12).

[Goo+12] Michael F. Goodchild, Huadong Guo, Alessandro Annoni, et al. “Next-generation Digital Earth”. In: *Proceedings of the National Academy of Sciences* 109.28 (2012), pp. 11088–11094. eprint: https://www.pnas.org/content/109/28/11088.full.pdf (cit. on p. 165).

[Gor99] Al Gore. “The Digital Earth: understanding our planet in the 21st century”. In: (1999) (cit. on pp. 31, 34).

[GR14] Clara Greed and Marion Roberts. *Introducing urban design: interventions and responses*. Routledge, 2014 (cit. on p. 5).

[Gra+17] Alejandro Graciano, Antonio J. Rueda, Lidia Ortega, and Francisco R. Feito. “Towards a Hybrid Framework for the Visualization and Analysis of 3D Spatial Data”. In: *Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics*. UrbanGIS’17. Redondo Beach, CA, USA: Association for Computing Machinery, 2017 (cit. on p. 65).

[Grö95] Eduard Gröller. “Nonlinear ray tracing: Visualizing strange worlds”. In: *The Visual Computer* 11.5 (1995), pp. 263–274 (cit. on p. 68).

[Gus+19] Dennis Guse, Henrique R. Orefice, Gabriel Reimers, and Oliver Hohlfeld. “TheFragebogen: A Web Browser-based Questionnaire Framework for Scientific Research”. In: *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. 2019, pp. 1–3 (cit. on p. 110).

[Häg70] Torsten Hägerstraand. “What about people in regional science?” In: *Papers in regional science* 24.1 (1970), pp. 7–24 (cit. on p. 36).

[Han07] Malgorzata Hanzl. “Information technology as a tool for public participation in urban planning: a review of experiments and potentials”. In: *Design studies* 28.3 (2007), pp. 289–307 (cit. on p. 6).

[Har+15] Tilo Hartmann, Werner Wirth, Peter Vorderer, et al. “Spatial presence theory: State of the art and challenges ahead”. In: *Immersed in Media*. Springer, 2015, pp. 115–135 (cit. on p. 23).

[Har+95] Beverly L. Harrison, Hiroshi Ishii, Kim J. Vicente, and William A. S. Buxton. “Transparent Layered User Interfaces: An Evaluation of a Display Design to Enhance Focused and Divided Attention”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’95. Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, pp. 317–324 (cit. on pp. 57, 58).

[HC17] Charlotte Hoarau and Sidonie Christophe. “Cartographic continuum rendering based on color and texture interpolation to enhance photo-realism perception”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 127 (2017), pp. 27–38 (cit. on p. 35).

[HC19] Wumeng Huang and Jing Chen. “A multi-scale VR navigation method for VR globes”. In: *International Journal of Digital Earth* 12.2 (2019), pp. 228–249. eprint: https://doi.org/10.1080/17538947.2018.1426646 (cit. on p. 165).

[HCM13] Charlotte Hoarau, Sidonie Christophe, and Sébastien Mustière. “Mixing, blending, merging or scrambling topographic maps and orthoimagery in geovisualizations”. In: *International Cartographic Conference*. 2013 (cit. on pp. 22, 56).

[HE12] Christopher Healey and James Enns. “Attention and visual memory in visualization and computer graphics”. In: *IEEE transactions on visualization and computer graphics* 18.7 (2012), pp. 1170–1188 (cit. on pp. 15, 35).

[Hea97] Patsy Healey. *Collaborative planning: Shaping places in fragmented societies*. Macmillan International Higher Education, 1997 (cit. on p. 6).

[Heg11] Mary Hegarty. "The cognitive science of visual-spatial displays: Implications for design". In: *Topics in cognitive science* 3.3 (2011), pp. 446–474 (cit. on p. 36).

[HF05] Jonathan M Holmes and Sherry L Fawcett. "Testing distance stereoacuity with the Frisby-Davis 2 (FD2) test". In: *American journal of ophthalmology* 139.1 (2005), pp. 193–195 (cit. on p. 20).

[HHL04] Timothy Hughes, Young Hyun, and David A. Liberles. "Visualising very large phylogenetic trees in three dimensional hyperbolic space". In: *BMC Bioinformatics* 5 (2004), pp. 1–6 (cit. on p. 79).

[HK20] Youngjib Ham and Jaeyoon Kim. "Participatory Sensing and Digital Twin City: Updating Virtual City Models for Enhanced Risk-Informed Decision-Making". In: *Journal of Management in Engineering* 36.3 (2020), p. 04020005 (cit. on p. 32).

[Hon+17] Sungsoo Hong, Rafal Kocielnik, Min-Joon Yoo, et al. "Designing interactive distance cartograms to support urban travelers". In: *2017 IEEE Pacific Visualization Symposium (PacificVis)*. 2017, pp. 81–90 (cit. on p. 164).

[How96] Richard J Howarth. "History of the stereographic projection and its early use in geology". In: *Terra Nova* 8.6 (1996), pp. 499–513 (cit. on p. 81).

[HS05] Geoffrey S Hubona and Gregory W Shirah. "Spatial cues in 3D visualization". In: *Ambient Intelligence for Scientific Discovery*. Springer, 2005, pp. 104–128 (cit. on p. 19).

[HT06] Marc Hassenzahl and Noam Tractinsky. "User experience-a research agenda". In: *Behaviour & information technology* 25.2 (2006), pp. 91–97 (cit. on p. 27).

[HTW90] Ping-Kang Hsiung, Robert H. Thibadeau, and Michael Wu. "T-Buffer: Fast Visualization of Relativistic Effects in Space-Time". In: 24.2 (Feb. 1990), pp. 83–88 (cit. on p. 68).

[HU88] Atsuki Higashiyama and Eizo Ueyama. "The perception of vertical and horizontal distances in outdoor settings". In: *Perception & Psychophysics* 44.2 (1988), pp. 151–156 (cit. on p. 20).

[Hur+18a] Jan Hurtado, Eduardo Albuquerque, Daniel Radetic, et al. "Floating Hamster Ball: A Locomotion Method for Free Flight in Virtual Environments". In: *2018 20th Symposium on Virtual and Augmented Reality (SVR)*. 2018, pp. 183–191 (cit. on pp. 80, 83, 149).

[Hur+18b] Christophe Hurter, Nathalie Henry Riche, Steven M Drucker, et al. "Fiberclay: Sculpting three dimensional trajectories to reveal structural insights". In: *IEEE transactions on visualization and computer graphics* 25.1 (2018), pp. 704–714 (cit. on p. 36).

[Hur05] Karel Hurts. "Common Region and Spatial Performance Using Map-Like Displays". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 49.17 (2005), pp. 1593–1597. eprint: https://doi.org/10.1177/154193120504901720 (cit. on p. 55).

[HV09] Danny Holten and Jarke J Van Wijk. "Force-directed edge bundling for graph visualization". In: *Computer graphics forum*. Vol. 28. 3. Wiley Online Library. 2009, pp. 983–990 (cit. on p. 36).

[Ign+19] M Ignatius, NH Wong, M Martin, and S Chen. “Virtual Singapore integration with energy simulation and canopy modelling for climate assessment”. In: *IOP Conference Series: Earth and Environmental Science*. Vol. 294. 1. IOP Publishing. 2019, p. 012018 (cit. on p. 32).

[Ing93] Tim Ingold. “Globes and spheres”. In: *Environmentalism: The view from anthropology* 32 (1993), p. 31 (cit. on p. 79).

[Ise+13] P. Isenberg, P. Dragicevic, W. Willett, A. Bezerianos, and J. Fekete. “Hybrid-Image Visualization for Large Viewing Environments”. In: *IEEE Transactions on Visualization & Computer Graphics* 19.12 (Dec. 2013), pp. 2346–2355 (cit. on p. 48).

[Ito+16] Takayuki Itoh, Kim Marriott, Falk Schreiber, and Uwe Wössner. “Immersive Analytics: A new multidisciplinary initiative to explore future interaction technologies for data analytics”. In: *Shonan Reports* (2016) (cit. on p. 50).

[Jac+09] R. Jacobs, S. Hawking, M. Cabbage, and L. Hawking. *Apollo: Through the Eyes of the Astronauts*. Harry N. Abrams, 2009 (cit. on p. 167).

[Jac14] Florence Jacquinod. “Production, pratique et usages des géovisualisations 3D dans l’aménagement du territoire”. PhD thesis. Université de St Etienne, 2014 (cit. on p. 75).

[JE12] W. Javed and N. Elmqvist. “Exploring the design space of composite visualization”. In: *2012 IEEE Pacific Visualization Symposium*. 2012, pp. 1–8 (cit. on pp. 57, 60).

[Jen+18] Bernhard Jenny, Daniel M Stephen, Ian Muehlenhaus, et al. “Design principles for origin-destination flow maps”. In: *Cartography and Geographic Information Science* 45.1 (2018), pp. 62–75 (cit. on p. 36).

[Jen12] Bernhard Jenny. “Adaptive Composite Map Projections”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (Dec. 2012), pp. 2575–2582 (cit. on p. 80).

[JGE12] Waqas Javed, Sohaib Ghani, and Niklas Elmqvist. “Polyzoom: Multiscale and Multi-focus Exploration in 2d Visual Spaces”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’12. Austin, Texas, USA: Association for Computing Machinery, 2012, pp. 287–296 (cit. on p. 58).

[Jon+12] Brett Jones, Rajinder Sodhi, David Forsyth, Brian Bailey, and Giuliano Maciocci. “Around device interaction for multiscale navigation”. In: *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services - MobileHCI ’12*. New York, New York, USA: ACM Press, 2012, p. 83 (cit. on p. 98).

[JS20] Davina Jackson and Richard Simpson. “Digital City: An Urban Perspective on Digital Earth”. In: *Manual of Digital Earth*. Springer, Singapore, 2020, pp. 527–563 (cit. on pp. 31, 33, 39, 40).

[KBÖ15] Hulusi Kafaligonul, Bruno G Breitmeyer, and Haluk Öğmen. “Feedforward and feedback processes in vision”. In: *Frontiers in psychology* 6 (2015), p. 279 (cit. on p. 9).

[KD09] SeungJun Kim and Anind K. Dey. “Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation”. In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, 2009, p. 133 (cit. on pp. 163, 164).

[Ken+93] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness". In: *The International Journal of Aviation Psychology* 3.3 (1993), pp. 203–220. eprint: https://doi.org/10.1207/s15327108ijap0303_3 (cit. on p. 122).

[Ken63] John F Kennedy. "We choose to go to the Moon". In: *Speech presented at Address at Rice University on the Nation's Space Effort in Rice University, Houston (1962, September 12)* (1963) (cit. on p. 168).

[KH13] Christopher C. M. Kyba and Franz Hölker. "Do artificially illuminated skies affect biodiversity in nocturnal landscapes?" In: *Landscape Ecology* 28.9 (Nov. 2013), pp. 1637–1640 (cit. on p. 101).

[KH20] Eva Kassens-Noor and Arend Hintze. "Cities of the Future? The Potential Impact of Artificial Intelligence". In: *AI* 1.2 (2020), pp. 192–197 (cit. on p. 32).

[Kir95] David Kirsh. "The intelligent use of space". In: *Artificial Intelligence* 73.1 (1995). Computational Research on Interaction and Agency, Part 2, pp. 31–68 (cit. on pp. 48, 56).

[KK19] A. Krekhov and J. Krüger. "Deadeye: A Novel Preattentive Visualization Technique Based on Dichoptic Presentation". In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 936–945 (cit. on p. 15).

[KL01] Frank L Kooi and Marcel Lucassen. "Visual comfort of binocular and 3D displays". In: *Human Vision and Electronic Imaging VI*. Vol. 4299. International Society for Optics and Photonics. 2001, pp. 586–592 (cit. on p. 20).

[KL04] Mei-Po Kwan and Jiyeong Lee. "Geovisualization of human activity patterns using 3D GIS: a time-geographic approach". In: *Spatially integrated social science* 27 (2004), pp. 721–744 (cit. on p. 30).

[Kof35] Kurt Koffka. *Principles of Gestalt Psychology*. Harcourt, Brace &Co, 1935 (cit. on p. 13).

[Kol95] Eugenia M Kolasinski. *Simulator sickness in virtual environments*. Vol. 1027. US Army Research Institute for the Behavioral and Social Sciences, 1995 (cit. on p. 148).

[Kra03] Menno-Jan Kraak. "Geovisualization illustrated". In: *ISPRS journal of photogrammetry and remote sensing* 57.5-6 (2003), pp. 390–399 (cit. on p. 34).

[Kra08] Menno-Jan Kraak. "Geovisualization and time–new opportunities for the space–time cube". In: *Geographic visualization: concepts, tools and applications* (2008), pp. 293–306 (cit. on p. 36).

[Kre+20] A. Krekhov, S. Cmentowski, A. Waschk, and J. Krüger. "Deadeye Visualization Revisited: Investigation of Preattentiveness and Applicability in Virtual Environments". In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (2020), pp. 547–557 (cit. on p. 15).

[KSR12] Kristian Kloeckl, Oliver Senn, and Carlo Ratti. "Enabling the real-time city: LIVE Singapore!" In: *Journal of Urban Technology* 19.2 (2012), pp. 89–112 (cit. on pp. 29, 30).

[Kur+14] Kuno Kurzhals, Brian Fisher, Michael Burch, and Daniel Weiskopf. "Evaluating visual analytics with eye tracking". In: *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*. ACM. 2014, pp. 61–69 (cit. on p. 22).

[Kur+16] Kuno Kurzhals, Brian Fisher, Michael Burch, and Daniel Weiskopf. “Eye tracking evaluation of visual analytics”. In: *Information Visualization* 15.4 (2016), pp. 340–358. eprint: https://doi.org/10.1177/1473871615609787 (cit. on p. 111).

[Kwo+16] Oh-Hyun Kwon, Chris Muelder, Kyungwon Lee, and Kwan-Liu Ma. “A Study of Layout, Rendering, and Interaction Methods for Immersive Graph Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 22.7 (2016), pp. 1802–1815 (cit. on pp. 80, 82).

[Lan+21] Fangfei Lan, Michael Young, Lauren Anderson, et al. “Visualization in Astrophysics: Developing New Methods, Discovering Our Universe, and Educating the Earth”. In: (May 2021). arXiv: 2106.00152 (cit. on p. 165).

[LAP17] María-Jesús Lobo, Caroline Appert, and Emmanuel Pietriga. “MapMosaic: dynamic layer compositing for interactive geovisualization”. In: *International Journal of Geographical Information Science* 31.9 (2017), pp. 1818–1845. eprint: https://doi.org/10.1080/13658816.2017.1325488 (cit. on pp. 56, 58, 59).

[Las89] Piotr H. Laskowski. “The Traditional and Modern Look at Tissot’s Indicatrix”. In: *The American Cartographer* 16.2 (1989), pp. 123–133. eprint: https://doi.org/10.1559/152304089783875497 (cit. on p. 137).

[LaV00] Joseph J LaViola Jr. “A discussion of cybersickness in virtual environments”. In: *ACM Sigchi Bulletin* 32.1 (2000), pp. 47–56 (cit. on p. 21).

[LCL13] Hui Lin, Min Chen, and Guonian Lu. “Virtual geographic environment: a workspace for computer-aided geographic experiments”. In: *Annals of the Association of American Geographers* 103.3 (2013), pp. 465–482 (cit. on p. 39).

[LD11] David Lloyd and Jason Dykes. “Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study”. In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 2498–2507 (cit. on p. 36).

[LG01] Hui Lin and Jianhua Gong. “Exploring virtual geographic environments”. In: *Geographic Information Sciences* 7.1 (2001), pp. 1–7 (cit. on p. 38).

[LH14] Zhicheng Liu and Jeffrey Heer. “The effects of interactive latency on exploratory visual analysis”. In: *IEEE transactions on visualization and computer graphics* 20.12 (2014), pp. 2122–2131 (cit. on p. 30).

[LH88] Margaret Livingstone and David Hubel. “Segregation of form, color, movement, and depth: anatomy, physiology, and perception”. In: *Science* 240.4853 (1988), pp. 740–749 (cit. on p. 14).

[Lin+13] Hui Lin, Min Chen, Guonian Lu, et al. “Virtual geographic environments (VGEs): a new generation of geographic analysis tool”. In: *Earth-Science Reviews* 126 (2013), pp. 74–84 (cit. on p. 39).

[Lin+17] Junguo Lin, Dewen Cheng, Cheng Yao, and Yongtian Wang. “Retinal projection head-mounted display”. In: *Frontiers of Optoelectronics* 10.1 (2017), pp. 1–8 (cit. on p. 21).

[Liu+20] Jiazhou Liu, Arnaud Prouzeau, Barrett Ens, and Tim Dwyer. “Design and Evaluation of Interactive Small Multiples Data Visualisation in Immersive Spaces”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2020, pp. 588–597 (cit. on pp. 50, 197).

[LM14a] Wei Luo and Alan M MacEachren. “Geo-social visual analytics”. In: *Journal of spatial information science* 2014.8 (2014), pp. 27–66 (cit. on pp. 29, 40).

[LM14b] Saturnino Luz and Masood Masoodian. “Readability of a Background Map Layer under a Semi-Transparent Foreground Layer”. In: *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. AVI ’14. Como, Italy: Association for Computing Machinery, 2014, pp. 161–168 (cit. on p. 58).

[LOD18] Lieselot Lapon, Kristien Ooms, and Philippe De Maeyer. “The Influence of the Web Mercator Projection on the Global-Scale Cognitive Map of Web Map Users”. In: *Lecture Notes in Geoinformation and Cartography*. Vol. 0. 198809. 2018, pp. 79–82 (cit. on p. 165).

[Lor+08] Haik Lorenz, Matthias Trapp, Jürgen Döllner, and Markus Jobst. “Interactive Multi-Perspective Views of Virtual 3D Landscape and City Models”. In: *The European Information Society: Taking Geoinformation Science One Step Further*. Ed. by Lars Bernard, Anders Friis-Christensen, and Hardy Pundt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 301–321 (cit. on pp. 71, 73, 83, 143, 157–161, 164, 184, 198).

[LPA15] María-Jesús Lobo, Emmanuel Pietriga, and Caroline Appert. “An Evaluation of Interactive Map Comparison Techniques”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2015, pp. 3573–3582 (cit. on pp. 57–60, 101).

[LS10] Z. Liu and J. Stasko. “Mental Models, Visual Reasoning and Interaction in Information Visualization: A Top-down Perspective”. In: *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), pp. 999–1008 (cit. on p. 55).

[LT92] Robert Laurini and Derek Thompson. *Fundamentals of spatial information systems*. Vol. 37. Academic press, 1992 (cit. on p. 52).

[Lu+18] Feiyu Lu, Difeng Yu, Hai-Ning Liang, et al. “Evaluating engagement level and analytical support of interactive visualizations in virtual reality environments”. In: *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2018, pp. 143–152 (cit. on p. 47).

[Lyn60] Kevin Lynch. *The image of the city*. Vol. 11. MIT press, 1960 (cit. on pp. 22, 66, 76).

[Mac11] Alan Maceachren. “The Roles of Maps, from Some Truth with Maps: A Primer on Symbolization and Design”. English (US). In: *The Map Reader*. John Wiley and Sons, Apr. 2011, pp. 244–251 (cit. on p. 59).

[Mah+18] T. Mahmood, E. Butler, N. Davis, J. Huang, and A. Lu. “Building Multiple Coordinated Spaces for Effective Immersive Analytics through Distributed Cognition”. In: *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*. 2018, pp. 1–11 (cit. on pp. 55, 61, 132).

[Mal19] Stephen Mallory. “To the Mun: Kerbal Space Program as Playful, Educational Experience”. In: *HCI in Games*. Ed. by Xiaowen Fang. Cham: Springer International Publishing, 2019, pp. 320–332 (cit. on p. 166).

[Man+18] Steve Mann, Tom Furness, Yu Yuan, Jay Iorio, and Zixin Wang. “All Reality: Virtual, Augmented, Mixed (X), Mediated (X, Y), and Multimediated Reality”. In: *arXiv* (2018), arXiv–1804 (cit. on p. 38).

[Mar+18a] Kim Marriott, Jian Chen, Marcel Hlawatsch, et al. “Immersive analytics: Time to reconsider the value of 3d for information visualisation”. In: *Immersive analytics*. Springer, 2018, pp. 25–55 (cit. on pp. 21, 27).

[Mar+18b] Kim Marriott, Jian Chen, Marcel Hlawatsch, et al. “Just 5 questions: Toward a design framework for immersive analytics”. In: *Immersive Analytics*. Springer, 2018, pp. 259–288 (cit. on p. 27).

[Mar+18c] Kim Marriott, Falk Schreiber, Tim Dwyer, et al. *Immersive Analytics*. Vol. 11190. Springer, 2018 (cit. on pp. 21, 23, 27).

[MAS15] Ali Mahdavi-Amiri, Troy Alderson, and Faramarz Samavati. “A Survey of Digital Earth”. In: *Computers & Graphics* 53 (2015), pp. 95–117 (cit. on p. 165).

[McC+18] Jon McCormack, Jonathan C Roberts, Benjamin Bach, et al. “Multisensory immersive analytics”. In: *Immersive Analytics*. Springer, 2018, pp. 57–94 (cit. on p. 27).

[MCC87] B. H. MCCORMICK. “Visualisation in Scientific Computing”. In: *Computer Graphics* 21 (1987), p. 6 (cit. on p. 25).

[Mcg+20] Mark Mcgill, Aidan Kehoe, Euan Freeman, and Stephen Brewster. “Expanding the Bounds of Seated Virtual Workspaces”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 27.3 (2020), pp. 1–40 (cit. on p. 50).

[McL03] Hilary McLellan. *Virtual Realities*. Wyatt Digital, 2003 (cit. on p. 10).

[Men+19] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. “A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments”. In: *Computer Graphics Forum* 38.1 (2019), pp. 21–45. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13390 (cit. on p. 80).

[Mer+17] L. Merino, M. Ghafari, C. Anslow, and O. Nierstrasz. “CityVR: Gameful Software Visualization”. In: *2017 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. 2017, pp. 633–637 (cit. on p. 47).

[MFW08] Eliana Medina, Ruth Fruland, and Suzanne Weghorst. “Virtusphere: Walking in a Human Size VR “Hamster Ball””. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 52.27 (2008), pp. 2102–2106. eprint: https://doi.org/10.1177/154193120805202704 (cit. on p. 80).

[MGV95] A David Milner, Melvyn A Goodale, and Algis J Vingrys. *The visual brain in action (Vol. 27)*. Oxford Psychology Series. Oxford University Press, Oxford, 1995 (cit. on p. 12).

[Min80] Marvin Minsky. “Telepresence”. In: *Omni* (1980) (cit. on p. 23).

[Mir+16] Fabio Miranda, Harish Doraiswamy, Marcos Lage, et al. “Urban pulse: Capturing the rhythm of cities”. In: *IEEE transactions on visualization and computer graphics* 23.1 (2016), pp. 791–800 (cit. on p. 30).

[Mir+18a] Fabio Miranda, Harish Doraiswamy, Marcos Lage, et al. “Shadow accrual maps: Efficient accumulation of city-scale shadows over time”. In: *IEEE transactions on visualization and computer graphics* 25.3 (2018), pp. 1559–1574 (cit. on p. 30).

[Mir+18b] Fabio Miranda, Marcos Lage, Harish Doraiswamy, et al. “Time lattice: A data structure for the interactive visual analysis of large time series”. In: *Computer Graphics Forum*. Vol. 37. 3. Wiley Online Library. 2018, pp. 23–35 (cit. on p. 30).

[Mit96] William John Mitchell. *City of bits: space, place, and the infobahn*. MIT press, 1996 (cit. on p. 31).

[MK01] Alan M. MacEachren and Menno-Jan Kraak. "Research Challenges in Geovisualization". In: *Cartography and Geographic Information Science* 28.1 (2001), pp. 3–12. eprint: https://doi.org/10.1559/152304001782173970 (cit. on p. 34).

[MK94] Paul Milgram and Fumio Kishino. "A taxonomy of mixed reality visual displays". In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), pp. 1321–1329 (cit. on pp. 37, 38).

[Mor14] David A Morrison. *The Book of Trees: Visualizing Branches of Knowledge.—By Manuel Lima. Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations.—By Isabel Meirelles*. Oxford University Press, 2014 (cit. on p. 13).

[Mös+08] Sebastian Möser, Patrick Degener, Roland Wahl, and Reinhard Klein. "Context Aware Terrain Visualization for Wayfinding and Navigation". In: *Computer Graphics Forum* 27.7 (2008), pp. 1853–1860. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2008.01332.x (cit. on pp. 71–73, 164).

[MRM84] Timothy P McNamara, Roger Ratcliff, and Gail McKoon. "The mental representation of knowledge acquired from maps." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10.4 (1984), p. 723 (cit. on p. 55).

[MT99] Mark Mon-Williams and James R Tresilian. "Some recent studies on the extraretinal contribution to distance perception". In: *Perception* 28.2 (1999), pp. 167–181 (cit. on p. 19).

[Mun00] Tamara Munzner. "Interactive Visualization of Large Graphs and Networks". PhD thesis. 2000, p. 149 (cit. on p. 79).

[Mun14] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014 (cit. on pp. 15, 16, 22, 26, 35).

[Mun98] Tamara Munzner. "Exploring large graphs in 3D hyperbolic space". In: *IEEE Computer Graphics and Applications* 18.4 (1998), pp. 18–23 (cit. on p. 79).

[MV17] Matthias Möller and Cornelis Vuik. "On the impact of quantum computing technology on future developments in high-performance scientific computing". In: *Ethics and Information Technology* 19.4 (2017), pp. 253–269 (cit. on p. 31).

[Nan+11] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. "Mid-air pan-and-zoom on wall-sized displays". In: *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. New York, New York, USA: ACM Press, 2011, p. 177 (cit. on p. 98).

[NC14] Jeanne Nakamura and Mihaly Csikszentmihalyi. "The concept of flow". In: *Flow and the foundations of positive psychology*. Springer, 2014, pp. 239–263 (cit. on p. 25).

[NDV12] Till Nagel, Erik Duval, and Andrew Vande Moere. "Interactive Exploration of Geospatial Network Visualization". In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '12. Austin, Texas, USA: Association for Computing Machinery, 2012, pp. 557–572 (cit. on p. 61).

[Neu49] Otto Neugebauer. "The Early History of the Astrolabe. Studies in Ancient Astronomy IX". In: *Isis* 40.3 (1949), pp. 240–256 (cit. on p. 81).

[NFF51] Hans Namuth, Paul Falkenberg, and Morton Freldman. *Jackson Pollock 51*. 1951 (cit. on p. 172).

[ODF15] Kristien Ooms, Philippe De Maeyer, and Veerle Fack. "Listen to the map user: cognition, memory, and expertise". In: *The Cartographic Journal* 52.1 (2015), pp. 3–19 (cit. on p. 36).

[OMS00] David O'Sullivan, Alastair Morrison, and John Shearer. "Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach". In: *International Journal of Geographical Information Science* 14.1 (Jan. 2000), pp. 85–104 (cit. on p. 164).

[Ort+16] Thomas Ortner, Johannes Sorger, Harald Steinlechner, et al. "Vis-a-ware: Integrating spatial and non-spatial visualization for visibility-aware urban planning". In: *IEEE Transactions on Visualization and Computer Graphics* 23.2 (2016), pp. 1139–1151 (cit. on pp. 22, 65).

[OT08] Heather L O'Brien and Elaine G Toms. "What is user engagement? A conceptual framework for defining user engagement with technology". In: *Journal of the American society for Information Science and Technology* 59.6 (2008), pp. 938–955 (cit. on p. 25).

[PA08] Emmanuel Pietriga and Caroline Appert. "Sigma Lenses: Focus-Context Transitions Combining Space, Time and Translucence". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: Association for Computing Machinery, 2008, pp. 1343–1352 (cit. on p. 87).

[PA18] RT Pramod and SP Arun. "Symmetric objects become special in perception because of generic computations in neurons". In: *Psychological science* 29.1 (2018), pp. 95–109 (cit. on p. 22).

[Pai+14] Jose Gustavo S Paiva, William Robson Schwartz, Helio Pedrini, and Rosane Minghim. "An approach to supporting incremental visual data classification". In: *IEEE transactions on visualization and computer graphics* 21.1 (2014), pp. 4–17 (cit. on p. 29).

[Pas+14] Sebastian Pasewaldt, Amir Semmo, Matthias Trapp, and Jürgen Döllner. "Multi-perspective 3D panoramas". In: *International Journal of Geographical Information Science* 28.10 (2014), pp. 2030–2051. eprint: https://doi.org/10.1080/13658816.2014.922686 (cit. on pp. 71, 73, 78, 157, 159–161, 164, 184).

[PD13] Sebastian Pasewaldt and Jürgen Döllner. "Interacting with Multi-Perspective Views: Challenges and Approaches". In: *Proceedings of the 1st ACM SIGSPATIAL International Workshop on MapInteraction*. MapInteract '13. Orlando, Florida: Association for Computing Machinery, 2013, pp. 44–47 (cit. on p. 71).

[Pei+19a] Etienne Peillard, Ferran Argelaguet, Jean-Marie Normand, Anatole Lécuyer, and Guillaume Moreau. "Studying Exocentric Distance Perception in Optical See-Through Augmented Reality". In: *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2019, pp. 115–122 (cit. on p. 21).

[Pei+19b] Etienne Peillard, Thomas Thebaud, Jean-Marie Norrnand, et al. "Virtual Objects Look Farther on the Sides: The Anisotropy of Distance Perception in Virtual Reality". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 227–236 (cit. on p. 21).

[Pei+20] Etienne Peillard, Yuta Itoh, Jean-Marie Normand, et al. “Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality?” In: *ISMAR 2020-19th IEEE International Symposium on Mixed and Augmented Reality*. 2020 (cit. on p. 21).

[Pep19] David Pepper. *The roots of modern environmentalism*. Vol. 6. Routledge, 2019 (cit. on p. 170).

[Pet11] Gregory A. Petsko. “The blue marble”. In: *Genome Biology* 12.4 (2011), p. 112 (cit. on p. 169).

[Pin+12] Cyprien Pindat, Emmanuel Pietriga, Olivier Chapuis, and Claude Puech. “JellyLens: content-aware adaptive lenses”. In: *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 2012, pp. 261–270 (cit. on pp. 35, 58, 68, 157).

[Piu+18] Thammathip Piumsomboon, Gun A. Lee, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. “Superman vs Giant: A Study on Spatial Perception for a Multi-Scale Mixed Reality Flying Telepresence Interface”. In: *IEEE Transactions on Visualization and Computer Graphics* 24.11 (Nov. 2018), pp. 2974–2982 (cit. on p. 182).

[PL17] Christopher Pettit and Simone Z. Leao. “Dashboard”. In: *Encyclopedia of Big Data*. Ed. by Laurie A. Schintler and Connie L. McNeely. Cham: Springer International Publishing, 2017, pp. 1–6 (cit. on p. 39).

[Pla+23] Plato, WRM Lamb, Robert Gregg Bury, Paul Shorey, and Harold North Fowler. *Plato in twelve volumes*. Heinemann, 1923 (cit. on pp. 168, 177).

[PLJ17] Christopher Pettit, Scott N Lieske, and Murad Jamal. “CityDash: Visualising a changing city using open data”. In: *International Conference on Computers in Urban Planning and Urban Management*. Springer. 2017, pp. 337–353 (cit. on p. 39).

[Pol+18] Nicholas Polys, Cecile Newcomb, Todd Schenk, Thomas Skuzinski, and Donna Dunay. “The value of 3D models and immersive technology in planning urban density”. In: *Proceedings of the 23rd International ACM Conference on 3D Web Technology*. 2018, pp. 1–4 (cit. on pp. 22, 66, 75).

[Poo08] Robert Poole. *Earthrise: How Man First Saw the Earth*. Newhaven and London: Yale University Press, Nov. 2008 (cit. on p. 170).

[Por+17] Thiago M Porcino, Esteban Clua, Daniela Trevisan, Cristina N Vasconcelos, and Luis Valente. “Minimizing cyber sickness in head mounted display systems: design guidelines and applications”. In: *2017 IEEE 5th international conference on serious games and applications for health (SeGAH)*. IEEE. 2017, pp. 1–6 (cit. on p. 21).

[PR94] Stephen Palmer and Irvin Rock. “Rethinking perceptual organization: The role of uniform connectedness”. In: *Psychonomic bulletin & review* 1.1 (1994), pp. 29–55 (cit. on p. 13).

[PTD11] Sebastian Pasewaldt, Matthias Trapp, and Jürgen Döllner. “Multiscale visualization of 3D geovirtual environments using view-dependent multi-perspective views”. In: *Journal of WSCG* 19.1-3 (2011), pp. 111–118 (cit. on pp. 71, 74).

[PTR17] Christopher Pettit, Andrew Tice, and Bill Randolph. “Using an online spatial analytics workbench for understanding housing affordability in Sydney”. In: *Seeing Cities Through Big Data*. Springer, 2017, pp. 233–255 (cit. on p. 39).

[PW06] Matthew D. Plumlee and Colin Ware. “Zooming versus Multiple Window Interfaces: Cognitive Costs of Visual Comparisons”. In: 13.2 (June 2006), pp. 179–209 (cit. on p. 58).

[Que+18] Anna Carolina Muller Queiroz, Amy M Kamarainen, Nicholas D Preston, and Maria Isabel da Silva Leme. “Immersive virtual environments and climate change engagement”. In: *iLRN 2018 Montana* (2018), p. 153 (cit. on p. 47).

[Ran+18] Nimesha Ranasinghe, Pravar Jain, Nguyen Thi Ngoc Tram, et al. “Season traveller: Multisensory narration for enhancing the virtual reality experience”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–13 (cit. on p. 24).

[Rau+15] Victoria Rautenbach, Serena Coetzee, Jochen Schiewe, and Arzu Cöltekin. “An assessment of visual variables for the cartographic design of 3D informal settlement models”. In: *27th International Cartographic Conference 16th General Assembly*. Proceedingstitel: Maps Connecting the World. Maps Connecting the World, Aug. 2015, online (cit. on p. 35).

[RB11] Paulo Raposo and Cynthia Ann Brewer. “Comparison of topographic map designs for overlay on orthoimage backgrounds”. In: *25th International Cartographic Conference*. 2011, pp. 3–8 (cit. on p. 58).

[Red+15] Khairi Reda, Andrew E. Johnson, Michael E. Papka, and Jason Leigh. “Effects of Display Size and Resolution on User Behavior and Insight Acquisition in Visual Exploration”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: Association for Computing Machinery, 2015, pp. 2759–2768 (cit. on p. 49).

[Rei+13] Christoph Reinhart, Timur Dogan, J Alstan Jakubiec, Tarek Rakha, and Andrew Sang. “Umi-an urban simulation environment for building energy use, daylighting and walkability”. In: *13th Conference of International Building Performance Simulation Association, Chambery, France*. Vol. 1. 2013 (cit. on p. 32).

[RFD20] P. Reipschläger, T. Flemisch, and R. Dachselt. “Personal Augmented Reality for Information Visualization on Large Interactive Displays”. In: *IEEE Transactions on Visualization and Computer Graphics* (2020), pp. 1–1 (cit. on p. 50).

[Ric70] Whitman Richards. “Stereopsis and stereoblindness”. In: *Experimental Brain Research* 10.4 (1970), pp. 380–388 (cit. on p. 20).

[Rie03] Bernhard E Riecke. *How Far Can We Get with Just Visual Information?: Path Integration and Spatial Updating Studies in Virtual Reality*. Logos Verlag, 2003 (cit. on p. 149).

[Riv09] Giuseppe Riva. “Is presence a technology issue? Some insights from cognitive sciences”. In: *Virtual reality* 13.3 (2009), pp. 159–169 (cit. on p. 24).

[RM93] George G Robertson and Jock D Mackinlay. “The document lens”. In: *Proceedings of the 6th annual ACM symposium on User interface software and technology*. 1993, pp. 101–108 (cit. on p. 35).

[RMC91] George G Robertson, Jock D Mackinlay, and Stuart K Card. “Cone trees: animated 3D visualizations of hierarchical information”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1991, pp. 189–194 (cit. on p. 26).

[RO16] Lisa Rebenitsch and Charles Owen. “Review on cybersickness in applications and visual displays”. In: *Virtual Reality* 20.2 (2016), pp. 101–125 (cit. on p. 21).

[Rob+08] George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, and John Stasko. "Effectiveness of animation in trend visualization". In: *IEEE transactions on visualization and computer graphics* 14.6 (2008), pp. 1325–1332 (cit. on p. 36).

[Rob+98] George Robertson, Mary Czerwinski, Kevin Larson, et al. "Data mountain: using spatial memory for document management". In: *Proceedings of the 11th annual ACM symposium on User interface software and technology*. 1998, pp. 153–162 (cit. on p. 26).

[Rob07] Jonathan C Roberts. "State of the art: Coordinated & multiple views in exploratory visualization". In: *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*. IEEE. 2007, pp. 61–71 (cit. on pp. 35, 60, 68).

[Ros67] Helen E Ross. "Water, fog and the size—distance invariance hypothesis". In: *British Journal of Psychology* 58.3-4 (1967), pp. 301–313 (cit. on p. 20).

[Rot+17] Robert E Roth, Arzu Çöltekin, Luciene Delazari, et al. "User studies in cartography: opportunities for empirical research on interactive maps and visualizations". In: *International Journal of Cartography* 3.sup1 (2017), pp. 61–89 (cit. on p. 36).

[Rot13] R. E. Roth. "An Empirically-Derived Taxonomy of Interaction Primitives for Interactive Cartography and Geovisualization". In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2356–2365 (cit. on p. 58).

[RR13] Joseph Ranalli and Jackie Ritzko. "Assessing the impact of video game based design projects in a first year engineering design course". In: *2013 IEEE Frontiers in Education Conference (FIE)*. 2013, pp. 530–534 (cit. on p. 166).

[RTC15] Kai-Florian Richter, Martin Tomko, and Arzu Cöltekin. "Are we there yet? Spatial cognitive engineering for situated human-computer interaction". In: (2015) (cit. on p. 36).

[Sal17] Rossella Salerno. "Mapping Urban Environment by Geometry(es) and Perception(s)". In: *Urban Design and Representation: A Multidisciplinary and Multisensory Approach*. Ed. by Barbara E.A. Piga and Rossella Salerno. Cham: Springer International Publishing, 2017, pp. 11–21 (cit. on pp. 66, 76).

[Sat+19] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, et al. "Augmented reality map navigation with freehand gestures". In: *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings* (2019), pp. 593–603 (cit. on p. 98).

[Sat+20] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Tobias Czauderna, and Bernhard Jenny. "Maps Around Me: 3D Multiview Layouts in Immersive Spaces". In: 4.ISS (Nov. 2020) (cit. on p. 50).

[SB20] Volkan Sevinc and Mehmet Ilker Berkman. "Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments". In: *Applied Ergonomics* 82 (2020), p. 102958 (cit. on p. 122).

[SB97] Andre Skupin and Barbara P. Buttenfield. "Spatial Metaphore for Visualizing Information Spaces". In: *AUTO-CARTO* (1997), pp. 116–125 (cit. on p. 35).

[SC02] William R. Sherman and Alan B. Craig. *Understanding Virtual Reality: Interface, Application, and Design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2002 (cit. on p. 38).

[Sch+11] Roeland Scheepens, Niels Willems, Huub Van de Wetering, et al. “Composite density maps for multivariate trajectories”. In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 2518–2527 (cit. on pp. 29, 30).

[Sch+13] Hans-Jörg Schulz, Thomas Nocke, Magnus Heitzler, and Heidrun Schumann. “A Design Space of Visualization Tasks”. In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2366–2375 (cit. on p. 100).

[Sch+15] Jörg Schaller, O Ertac, Sebastian Freller, Cristina Mattos, and Zeljko Rajcevic. “Geodesign apps and 3D modelling with CityEngine for the city of tomorrow”. In: *Digital Landscape Architecture* (2015), pp. 59–70 (cit. on p. 32).

[Sch10] Ulrike Schultze. “Embodiment and presence in virtual worlds: a review”. In: *Journal of Information Technology* 25.4 (2010), pp. 434–449 (cit. on p. 24).

[SCP95] Richard Stoakley, Matthew J. Conway, and Randy Pausch. “Virtual reality on a WIM”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’95*. Vol. 1. figure 1. New York, New York, USA: ACM Press, 1995, pp. 265–272 (cit. on p. 185).

[Sen+18] Neelava Sengupta, Josafath Israel Espinosa Ramos, Enmei Tu, et al. “From von Neumann Architecture and Atanasoffs ABC to Neuro-Morphic Computation and Kasabov’s NeuCube: Principles and Implementations”. In: *Learning Systems: From Theory to Practice*. Springer, 2018, pp. 1–28 (cit. on p. 31).

[SES18] Gurjot Singh, Stephen R Ellis, and J Edward Swan II. “The effect of focal distance, age, and brightness on near-field augmented reality depth matching”. In: *IEEE transactions on visualization and computer graphics* (2018) (cit. on p. 21).

[SG00] Dario D. Salvucci and Joseph H. Goldberg. “Identifying Fixations and Saccades in Eye-Tracking Protocols”. In: *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. ETRA ’00. Palm Beach Gardens, Florida, USA: Association for Computing Machinery, 2000, pp. 71–78 (cit. on p. 111).

[She+71] William E Shenk, Hugh Powell, Vincent V Salomonson, and William R Bandeen. “Meteorological uses of the stereographic horizon map projection”. In: *Journal of Applied Meteorology* 10.3 (1971), pp. 582–589 (cit. on p. 81).

[Sho92] Ken Shoemake. “ARCBALL: A user interface for specifying three-dimensional orientation using a mouse”. In: *Proceedings of Graphics Interface ’92*. Vol. 92. 1992, pp. 151–156 (cit. on p. 79).

[Sic+19] Ronell Sicat, Jiabao Li, Junyoung Choi, et al. “DXR: A Toolkit for Building Immersive Data Visualizations”. In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 715–725 (cit. on p. 95).

[SK17] Shun Sekiguchi and Takashi Komuro. “VolGrab: Realizing 3D View Navigationby Aerial Hand Gestures”. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. New York, NY, USA: ACM, May 2017, pp. 2064–2069 (cit. on p. 98).

[Sla09] Mel Slater. “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1535 (2009), pp. 3549–3557 (cit. on p. 24).

[Sla18] Mel Slater. “Immersion and the illusion of presence in virtual reality”. In: *British Journal of Psychology* 109.3 (2018), pp. 431–433 (cit. on p. 24).

[Sla99] Mel Slater. “Measuring presence: A response to the Witmer and Singer presence questionnaire”. In: *Presence* 8.5 (1999), pp. 560–565 (cit. on p. 24).

[SM15] Mar Santamaria-Varas and Pablo Martinez-Diez. “atNight: nocturnal landscapes and invisible networks”. In: *New Challenges for Data Design*. Springer, 2015, pp. 119–137 (cit. on p. 33).

[Sme] G Smets. “Designing for telepresence: The Delft virtual window system”. In: *Local applications of the ecological approach to human-machine systems* 2 (), pp. 182–207 (cit. on p. 24).

[Sny87] John Parr Snyder. *Map projections–A working manual*. Vol. 1395. US Government Printing Office, 1987 (cit. on pp. 81, 87).

[Sob16] Vivian Sobchack. “From Screen-Scape to Screen-Sphere: A Meditation in Medias Res”. In: *Screens: From Materiality to Spectatorship—A Historical and Theoretical Reassessment* (2016), pp. 157–175 (cit. on p. 80).

[SP08] Matthias Schonlau and Ellen Peters. “Graph comprehension: An experiment in displaying data as bar charts, pie charts and tables with and without the gratuitous 3rd dimension”. In: *Social Science Research Network Working Paper Series* (2008), pp. 1–16 (cit. on p. 26).

[SPG12] R San José, JL Pérez, and RM González. “Advances in 3D visualization of air quality data”. In: *Usage, Usability, and Utility of 3D City Models–European COST Action TU0801* (2012), p. 02002 (cit. on p. 22).

[Spi99] Lothar Spillmann. “From elements to perception: Local and global processing in visual neurons”. In: *Perception* 28.12 (1999), pp. 1461–1492 (cit. on p. 9).

[Spr+97] T. C. Sprenger, M. H. Gross, A. Eggenberger, and M. Kaufmann. “A Framework for Physically-Based Information Visualization”. In: *Visualization in Scientific Computing ’97*. Ed. by Wilfrid Lefer and Michel Grave. Vienna: Springer Vienna, 1997, pp. 71–83 (cit. on p. 79).

[SQR19] Ekaterina R. Stepanova, Denise Quesnel, and Bernhard E. Riecke. “Space—A Virtual Frontier: How to Design and Evaluate a Virtual Reality Experience of the Overview Effect”. In: *Frontiers in Digital Humanities* 6.April (Apr. 2019), pp. 1–22 (cit. on p. 170).

[SS05] Maria V Sanchez-Vives and Mel Slater. “From presence to consciousness through virtual reality”. In: *Nature Reviews Neuroscience* 6.4 (2005), pp. 332–339 (cit. on p. 24).

[SSF17] Alex D Singleton, Seth Spielman, and David Folch. *Urban analytics*. Sage, 2017 (cit. on p. 28).

[SSM13] Mirko Suznjevic, Lea Skorin-Kapov, and Maja Matijasevic. “The impact of user, system, and context factors on gaming QoE: A case study involving MMORPGs”. In: *2013 12th Annual Workshop on Network and Systems Support for Games (NetGames)*. 2013, pp. 1–6 (cit. on p. 97).

[Ste17] Stanley Smith Stevens. *Psychophysics: Introduction to its perceptual, neural and social prospects*. Routledge, 2017 (cit. on p. 18).

[SUC95] Mel Slater, Martin Usoh, and Yiorgos Chrysanthou. “The influence of dynamic shadows on presence in immersive virtual environments”. In: *Virtual environments’ 95*. Springer, 1995, pp. 8–21 (cit. on p. 24).

[Sut68] Ivan E Sutherland. “A head-mounted three dimensional display”. In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. 1968, pp. 757–764 (cit. on p. 26).

[SW97] Mel Slater and Sylvia Wilbur. “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments”. In: *Presence: Teleoperators & Virtual Environments* 6.6 (1997), pp. 603–616 (cit. on pp. 23–25).

[TD19] M. Trapp and J. Döllner. “Interactive Close-Up Rendering for Detail+Overview Visualization of 3D Digital Terrain Models”. In: *2019 23rd International Conference Information Visualisation (IV)*. 2019, pp. 275–280 (cit. on pp. 54, 58, 59).

[TGB90] Edward R Tufte, Nora Hillman Goeler, and Richard Benson. *Envisioning information*. Vol. 126. Graphics press Cheshire, CT, 1990 (cit. on p. 25).

[THF06] Emine Mine Thompson, Margaret Horne, and David Fleming. “Virtual reality urban modelling-an overview”. In: *CONVR2006: 6th Conference of Construction Applications of Virtual Reality* (2006) (cit. on p. 28).

[Tho+18] Bruce H Thomas, Gregory F Welch, Pierre Dragicevic, et al. “Situated Analytics.” In: *Immersive analytics* 11190 (2018), pp. 185–220 (cit. on p. 27).

[Tom+12] Christian Tominski, Heidrun Schumann, Gennady Andrienko, and Natalia Andrienko. “Stacking-based visualization of trajectory attribute data”. In: *IEEE Transactions on visualization and Computer Graphics* 18.12 (2012), pp. 2565–2574 (cit. on pp. 30, 36).

[Tom+17] Christian Tominski, Stefan Gladisch, Ulrike Kister, Raimund Dachselt, and Heidrun Schumann. “Interactive lenses for visualization: An extended survey”. In: *Computer Graphics Forum*. Vol. 36. 6. Wiley Online Library. 2017, pp. 173–200 (cit. on p. 35).

[Tor15] Paul M Torrens. “Slipstreaming human geosimulation in virtual geographic environments”. In: *Annals of GIS* 21.4 (2015), pp. 325–344 (cit. on p. 38).

[TS85] Anne Treisman and Janet Souther. “Search asymmetry: A diagnostic for preattentive processing of separable features.” In: *Journal of Experimental Psychology: General* 114.3 (1985), p. 285 (cit. on p. 14).

[Tsa+16] Tsai-Hsuan Tsai, Hsien-Tsung Chang, Ming-Chun Yu, et al. “Design of a mobile augmented reality application: An example of demonstrated usability”. In: *International Conference on Universal Access in Human-Computer Interaction*. Springer. 2016, pp. 198–205 (cit. on p. 37).

[Tuk77] John W Tukey. *Exploratory data analysis*. Vol. 2. Reading, MA, 1977 (cit. on p. 25).

[Tur10] Fred Turner. *From counterculture to cyberculture: Stewart Brand, the Whole Earth Network, and the rise of digital utopianism*. University of Chicago Press, 2010 (cit. on p. 170).

[Uni20] Unity Technologies. *Unity 3D*. Apr. 20, 2020 (cit. on p. 95).

[US 01] Ann Arbor MI U.S. Environmental Protection Agency. *EPA's New Generation Mobile Source Emissions Model: Initial Proposal and Issues*. Apr. 2001 (cit. on p. 34).

[VC01] Scott Vallance and Paul Calder. “Multi-Perspective Images for Visualisation”. In: *Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing - Volume 11*. VIP '01. Sydney, Australia: Australian Computer Society, Inc., 2001, pp. 69–76 (cit. on pp. 73, 74, 76, 78, 79, 83, 157, 158, 160).

[VDV14] Silviu Vert, Bogdan Dragulescu, and Radu Vasiu. “LOD4AR: Exploring Linked Open Data with a Mobile Augmented Reality Web Application.” In: *International Semantic Web Conference (Posters & Demos)*. Citeseer. 2014, pp. 185–188 (cit. on p. 37).

[Vea+12] E. Veas, R. Grasset, E. Kruijff, and D. Schmalstieg. “Extended Overview Techniques for Outdoor Augmented Reality”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.4 (2012), pp. 565–572 (cit. on pp. 71–73, 76, 164).

[VGB05] Dhanraj Vishwanath, Ahna R Girshick, and Martin S Banks. “Why pictures look right when viewed from the wrong place”. In: *Nature neuroscience* 8.10 (2005), pp. 1401–1410 (cit. on pp. 18, 21, 61, 83, 154).

[Vis14] Dhanraj Vishwanath. “Toward a new theory of stereopsis.” In: *Psychological review* 121.2 (2014), p. 151 (cit. on p. 20).

[VKB14] Stefano Vezzani, Peter Kramer, and Paola Bressan. *Stereokinetic Effect, Kinetic Depth Effect, and Structure from Motion*. Ed. by Johan Wagemans. Oxford University Press, Nov. 2014, pp. 1–26 (cit. on p. 86).

[Wag+12] Johan Wagemans, James H Elder, Michael Kubovy, et al. “A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization.” In: *Psychological bulletin* 138.6 (2012), p. 1172 (cit. on p. 13).

[Wal15] Peter Walker. “Depicting visual motion in still images: forward leaning and a left to right bias for lateral movement”. In: *Perception* 44.2 (2015), pp. 111–128 (cit. on p. 17).

[Wan+19] L. Wang, J. Wu, X. Yang, and V. Popescu. “VR Exploration Assistance through Automatic Occlusion Removal”. In: *IEEE Transactions on Visualization and Computer Graphics* 25.5 (2019), pp. 2083–2092 (cit. on p. 69).

[War13] Colin Ware. *Information Visualization: Perception for Design*. Elsevier, 2013 (cit. on pp. 11, 13, 19, 20, 189).

[War99] Colin Ware. *Perception for Design*. San Diego, Calif.: Morgan Kaufmann, 1999 (cit. on p. 10).

[Wat+98] Benjamin Watson, Neff Walker, William Ribarsky, and Victoria Spaulding. “Effects of variation in system responsiveness on user performance in virtual environments”. In: *Human Factors* 40.3 (1998), pp. 403–414 (cit. on p. 20).

[Whi14] Frank White. *The Overview Effect: Space Exploration and Human Evolution, Third Edition*. Reston ,VA: American Institute of Aeronautics and Astronautics, Jan. 2014 (cit. on pp. 170, 171).

[Wid+15] Ivo Widjaja, Patrizia Russo, Chris Pettit, Richard Sinnott, and Martin Tomko. “Modeling coordinated multiple views of heterogeneous data cubes for urban visual analytics”. In: *International Journal of Digital Earth* 8.7 (2015), pp. 558–578. eprint: https://doi.org/10.1080/17538947.2014.942713 (cit. on p. 39).

[WKB19] Séamas Weech, Sophie Kenny, and Michael Barnett-Cowan. “Presence and cybersickness in virtual reality are negatively related: a review”. In: *Frontiers in psychology* 10 (2019), p. 158 (cit. on p. 21).

[WP16] Meng-Lin Wu and Voicu Popescu. “Multiperspective Focus+Context Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 22.5 (May 2016), pp. 1555–1567 (cit. on pp. 69, 157).

[WP18] M. Wu and V. Popescu. “Efficient VR and AR Navigation Through Multiperspective Occlusion Management”. In: *IEEE Transactions on Visualization and Computer Graphics* 24.12 (2018), pp. 3069–3080 (cit. on p. 69).

[WS98] Bob G Witmer and Michael J Singer. “Measuring presence in virtual environments: A presence questionnaire”. In: *Presence* 7.3 (1998), pp. 225–240 (cit. on p. 24).

[WSD11] Jo Wood, Aidan Slingsby, and Jason Dykes. “Visualizing the dynamics of London’s bicycle-hire scheme”. In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 46.4 (2011), pp. 239–251 (cit. on p. 36).

[Wue12] Donald J. Wuebbles. “Celebrating the “Blue Marble””. In: *Eos, Transactions American Geophysical Union* 93.49 (Dec. 2012), pp. 509–510 (cit. on p. 170).

[Yad+16] David B. Yaden, Jonathan Iwry, Kelley J. Slack, et al. “The overview effect: Awe and self-transcendent experience in space flight.” In: *Psychology of Consciousness: Theory, Research, and Practice* 3.1 (Mar. 2016), pp. 1–11 (cit. on p. 170).

[Yan+18a] Yalong Yang, Tim Dwyer, Bernhard Jenny, et al. “Origin-destination flow maps in immersive environments”. In: *IEEE transactions on visualization and computer graphics* 25.1 (2018), pp. 693–703 (cit. on p. 36).

[Yan+18b] Yalong Yang, Bernhard Jenny, Tim Dwyer, et al. “Maps and Globes in Virtual Reality”. In: *Computer Graphics Forum* 37.3 (2018), pp. 427–438. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13431 (cit. on pp. 79, 90, 165).

[Yan+19] Yan Yan, Ke Chen, Yu Xie, Yiming Song, and Yonghong Liu. “The Effects of Weight on Comfort of Virtual Reality Devices”. In: *Advances in Ergonomics in Design*. Ed. by Francisco Rebelo and Marcelo M. Soares. Cham: Springer International Publishing, 2019, pp. 239–248 (cit. on p. 130).

[YCB05] Yonggao Yang, J. X. Chen, and M. Beheshti. “Nonlinear perspective projections and magic lenses: 3D view deformation”. In: *IEEE Computer Graphics and Applications* 25.1 (2005), pp. 76–84 (cit. on p. 68).

[Yig+20] Tan Yigitcanlar, Kevin C Desouza, Luke Butler, and Farnoosh Roozkhosh. “Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature”. In: *Energies* 13.6 (2020), p. 1473 (cit. on p. 32).

[You+92] James W Youdas, Tom R Garrett, Vera J Suman, et al. “Normal Range of Motion of the Cervical Spine: An Initial Goniometric Study”. In: *Physical Therapy* 72.11 (Nov. 1992), pp. 770–780. eprint: https://academic.oup.com/ptj/article-pdf/72/11/770/10759768/ptj0770.pdf (cit. on p. 82).

[YPL07] Perry Pei-Ju Yang, Simon Yunuar Putra, and Wenjing Li. “Viewsphere: a GIS-based 3D visibility analysis for urban design evaluation”. In: *Environment and Planning B: Planning and Design* 34.6 (2007), pp. 971–992 (cit. on p. 22).

[Zha+12] X. Zhao, W. Zeng, X. D. Gu, et al. “Conformal Magnifier: A Focus+Context Technique with Local Shape Preservation”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.11 (2012), pp. 1928–1941 (cit. on pp. 68, 157).

[Zhe+14] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. “Urban computing: concepts, methodologies, and applications”. In: *ACM Transactions on Intelligent Systems and Technology (TIST)* 5.3 (2014), pp. 1–55 (cit. on p. 29).

[Zhe+16] Yixian Zheng, Wenchao Wu, Yuanzhe Chen, Huamin Qu, and Lionel M Ni. “Visual analytics in urban computing: An overview”. In: *IEEE Transactions on Big Data* 2.3 (2016), pp. 276–296 (cit. on pp. 6, 7, 29, 40).

[Zie18] Karolina M. Zielinska-Dabkowska. “Make lighting healthier”. In: *Nature* 553.7688 (Jan. 2018), pp. 274–276 (cit. on p. 101).

[ZKF03] Charles M Zaroff, Magosha Knutelska, and Thomas E Frumkes. “Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender”. In: *Investigative ophthalmology & visual science* 44.2 (2003), pp. 891–900 (cit. on p. 20).

[ZN94] Jiaje Zhang and Donald A. Norman. “Representations in Distributed Cognitive Tasks”. In: *Cognitive Science* 18.1 (1994), pp. 87–122. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog1801_3 (cit. on p. 55).

Websites

[App21] Apple. *WWDC 2021 Keynote*. 2021. URL: https://www.apple.com/apple-events/june-2021/ (visited on June 7, 2021) (cit. on pp. 163, 165).

[Auta] Singapore Land Authority. *Singapore Geospatial Master Plan*. URL: https://www.geospatial.sg/initiatives#masterplan (visited on Nov. 1, 2020) (cit. on p. 32).

[Autb] Singapore Land Authority. *Virtual Singapore*. URL: https://www.geospatial.sg/gw/virtual-singapore (visited on Nov. 1, 2020) (cit. on p. 32).

[Bis20] Bob Bishop. *Integrative modeling, simulation, visualization and analysis to better understand, predict and protect our planet*. Nov. 2020. URL: http://www.icesfoundation.org/Pages/Home.aspx (visited on Nov. 4, 2020) (cit. on p. 31).

[Fou20] OpenStreetMap Foundation. *OpenStreetMap*. 2020. URL: https://www.openstreetmap.org/ (visited on May 5, 2021) (cit. on p. 97).

[Gre] City of Greater Bendigo. *Greening Greater Bendigo*. URL: https://www.bendigo.vic.gov.au/Services/Trees-and-reserves/Greening-Greater-Bendigo (visited on Nov. 1, 2020) (cit. on p. 33).

[IGN20] IGN. *BD TOPO*. 2020. URL: https://geoservices.ign.fr/documentation/diffusion/telechargement-donnees-libres.html#bd-topo (visited on May 4, 2021) (cit. on pp. 97, 151).

[Map20] Mapbox. *Mapbox*. 2020. URL: https://www.mapbox.com/ (visited on Apr. 20, 2020) (cit. on pp. 96, 97, 151).

[Mét20] Nantes Métropole. *Nantes Portail Open Data*. 2020. URL: https://data.nantesmetropole.fr/ (visited on May 5, 2021) (cit. on p. 97).

[Mun17] Tamara Munzner. *Data Visualization as a Driver for Visual Cognition Research*. 2017. URL: http://www.cs.ubc.ca/~tmm/talks.html#opam17 (visited on May 4, 2018) (cit. on p. 15).

[Tec] Varjo Technologies. *Varjo XR-3*. URL: https://varjo.com/products/xr-3/ (visited on Jan. 7, 2021) (cit. on p. 49).

List of Figures

2.1 A three-stage model of visual information processing *(adapted from [War13])*. 11

2.2 Ranking of visual channels on their effectiveness, as proposed by Tamara Munzner [Mun14]. 16

2.3 Stevens' Power Law, showing the relations between physical intensity of a stimulus and its perceived intensity. 18

3.1 The different steps and phases in a typical urban development process, and the level to which different actors are involved in them. 46

4.1 A Google image search for the query "geographic information systems;" highlighting the majority of results which show multiple data layers illustrated as a vertical stack. 51

4.2 A vertical stack of layers of the University of Ottawa's campus master plan. . 52

4.3 The concept and implementation of Brooks and Whalley's 3D GIS layering system [BW05; BW08]. 53

4.4 Examples of Trapp and Döllner's rendering technique from [TD19]. 54

4.5 View of the vertical separation of streets and points of interest from the urban base map, as investigated in [ED15]. 54

4.6 Examples of the Multiple Coordinated Spaces from [Mah+18]. 55

4.7 Layers "A" and "B" composited using different strategies. Not shown: *blitting*, or *temporal multiplexing*; where the layers are interactively switched in place. 57

4.8 Example workflow of creating a composite map from different layers using Lobo et al.'s MapMosaic [LAP17] 59

4.9 Side views of 4 map layers arranged in a 2×2 grid (a), stacked vertically in a flat arrangement (b), and in a dynamic stack that tracks the viewer's position (c). 62

5.1 Mockups of the minimap (a) and lens (b) variants of overview+detail, and the fisheye variant of focus+context, as defined by [CKB09]. 67

5.2 Disocclusion of regions of interest through multiperspective cameras by Wu and Popescu [WP18]. 69

5.3 Ordinary (left) and panoramic (right) perspective views of a city scene. The information beside and behind the virtual camera is lost in the ordinary perspective view, whereas it is preserved in the panoramic perspective view. In both perspective mappings, the information far away from the viewpoint is compressed into to a singular line, the horizon (from [Böt+08]). 69
5.4 The characteristics of the complex logarithmic transformation Böttger et al. employ to provide a focus+context view, and its application to show a city block in global context by using a planar projection of the whole Earth as its input (adapted from [Böt+08]). 70
5.5 Front and side views of the geometry-bending approach to focus+context, with different configurations (from [Pas+14]) 71
5.6 Multiperspective focus+context view with local deformations of regions of interest, and disocclusion of façades (from [Mös+08]). 72
5.7 The "extended overview" approach by Veas et al. to overlay a secondary camera's context view over an AR view (from [Vea+12]). 72
5.8 A detail and overview of a maze (a), and Vallance and Calder's omnidirectional geometry bending approach to reveal context (b): "Data is distorted according to its distance from the user, so that close data is the same as a normal first person view, while distorted data approximates looking at that data from a second person viewpoint above the maze" (from [VC01]). 73
5.9 The simple parametrization of Vallance and Calder's [VC01] geometry bending technique (a), and the more complex b-spline (b) used by Pasewaldt et al. [PTD11]. 74
5.10 Four iterations on the layered data approach from section 4.4: (a) the "sky layer," providing an overview over a first-person view, (b) the "dome," bending the layer down to establish a connection with the ground, (c) the "bowl," an inversion of the dome, creating a focus+context view, and (d) a flattened bowl, creating a larger focus area where the additional layer is directly overlaid on the ground. Viewing angles of distant objects in the additional layer are shown with $\alpha_{1..4}$ in comparison with the equivalent original object's β. . . . 77
5.11 Stereographic projection of points on a spherical surface to secant or tangent planes. Images on parallel planes differ only by an isometric scaling factor. . 81
5.12 Adaptation of the inverse stereographic principle to project any point P in 3D space to a point P' on the surface of a sphere (a), and in an application for urban focus+context (b). The sphere intersects the digital city model as it is elevated by h. Projected and original viewing angles α and β are compared as in Figure 5.10. 84
5.13 Additional distortions in the inverse stereographic projection caused by uneven terrain: toward-facing slopes are expanded, while those facing away from the projection origin get compressed. Sufficiently high peaks can cause an inversion, where the away-facing slope and plains behind it may be occluded, or appear inverted if projected transparently. 85

5.14 The effects of flattening terrain outside the sphere for conformal stereographic projection (a), and a prevention of these possible gaps and occlusions by gradually flattening the terrain in a transition zone outside the sphere (b). 86
5.15 The construction of conical (a) and cylindrical (b) projection surfaces from an "equivalent" sphere, by placing tangential surfaces at the intersection between sphere and the ground plane. Any point projected on the cone or cylinder (shown on the right half) is lower than on its equivalent sphere (shown on the left half), as indicated by dashed horizontal lines. 88
5.16 Two variations of the projection origin's height shown for the cone (blue), compared to the original on the equivalent sphere's azimuth (red). If this height is increased, the projected context just behind the transition is stretched longer than it is originally. 89
5.17 Using the stereographic projection first on the surface of a whole planetary body (the source) to an intermediary plane, then a second time on the virtual sphere focus+context projection. The uniform features on the source body (blue) are used to illustrate the effects of each projection regarding expansion and compression: while the final projection is conformal to the source, it is non-uniformally compressed if its radius is smaller than the radius of the source body, or expanded if larger. 90

6.1 One of the two (identical) controllers of the HTC Vive set. It is tracked in translation and rotation, and has three main input methods: the trigger (magenta), the side or "grip" buttons (blue), and the touchpad (green). The trigger outputs its continuous depression level, and the grip buttons output a binary value if any side is depressed. The touchpad registers the x–y position of a finger on it and can output a binary state when it is clicked. 96
6.2 A bounding box (grey), which defines a frame within which map tiles (green) are displayed (blue). The user will only see the part of the map that intersects this box. 98
6.3 Different viewing angles demonstrated for the MCV stack. The center of each layer is at the same distance to the viewer, the arc between layers is the same. Rays from the camera to the layers are colored by their length relative to the central ray. White lines show the normals the viewing angles are referenced to. 99
6.4 Clockwise from top left: light pollution aerial orthophoto, nightly transportation network, nocturnal points of interest, and a heatmap of energy consumption by public lighting. 103
6.5 The configurations for the blitting (left), grid (middle), and stack (right) MCV systems as used for the user study. Marker placement is shown for the grid and stack — note the connecting line in the latter's case. 106
6.6 The interface for selecting and confirming candidate areas. 107
6.7 Histograms showing the demographics and habits of the 26 participants. Participant 1 and 21 highlighted in green and blue, respectively. 107

6.8 Histograms showing the demographics and experiences (measured by visual analog scale, 0–100) of the 26 participants. Participant 1 and 21 highlighted in green and blue, respectively. 108

6.9 **(a)** How often (in percent) each MCV system was ranked *first, second,* or *last* in terms of readability (legibility of the map layers), ease of use, and visual design by the users; **(b)** Pairwise comparisons scored with the Bradley-Terry model. 112

6.10 **(a)** Which proportion of participants gave the MCV systems a number of zero to three *first* or *last* rankings across all aspects (readability, ease of use, and visual design); **(b)** Which proportion gave the most *firsts* (and therefore *preferred*) or *lasts* (and therefore *disliked*) to each system in the pairwise comparisons. 113

6.11 **Left:** the number of participants who play 3D-based video games at different frequencies, and their distribution among those who *preferred* and *disliked* each system, according to the measurement shown in Figure 6.10; **Right:** the same for the frequency with which participants consult city maps. 114

6.12 Performance measurements for each MCV system (blit/grid/stack): scenario completion times (**a**), and the number of times gaze has shifted from one map to another per scenario (**b**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3). 115

6.13 Map navigation measurements for each MCV system (blit/grid/stack): how much they panned the map per scenario (**a**), and how much they zoomed in/out of the map (**b**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3). . . 116

6.14 Average saccade amplitude measurements for each MCV system (blit/grid/stack): split by the eye (or viewport) component (**a**), the head movement (or HMD rotation) component (**b**), and taken together to determine the overall gaze (**c**). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3). 117

6.15 An example of a polar plot, showing the density probability distributions of absolute saccades—the gaze, as projected on a sphere—in degrees, for each MCV system (blit/grid/stack) and subgrouped by the participant's frequency of video game play. 119

6.16 Further visuo-motor behavior measurements for each MCV system (blit/grid/stack): the average fixation durations between saccades (**a**), and the average peak velocities reached during saccades (**b**) (upper confidence intervals cut to increase legibility of smaller values). Further subgrouped by the participants' preference for an MCV system regarding its legibility (1), the frequency with which they play video games (2), and the frequency with which they consult city maps (3). 120

6.17 An overview over the 12 scenarios (color-coded) and the distribution of choices participants made between the three candidates in each. 123

6.18 Charts showing data for each scenario: gaze information, candidate choices, and layer mentions in explanations. Scenario 5 and 12 highlighted. 124

6.19 Clockwise from top left: light pollution aerial orthophoto, nocturnal transport network, points of interest, and energy consumption heatmap layers of scenario 5, with candidate area outlines. 125

6.20 Clockwise from top left: light pollution aerial orthophoto, nocturnal transport network, points of interest, and energy consumption heatmap layers of scenario 12, with candidate area outlines. 126

6.21 Overview of the gaze and explanations of each participant. Participant 1 and 21 highlighted. 127

6.22 Analysis of the two selected participants (1 and 21), showing differences in choices, explanations, and visual attention as measured by time spent gazing at different layers. 128

7.1 Inverse stereographic projection of a point P to P' on a sphere with a radius of $\frac{1}{2}$, and its center at $(0,0,\frac{1}{2})$. 136

7.2 External and internal views of a simple inverse stereographic projection of a flat, textured plane, with the original view superimposed on half of the internal view for comparison. The unchanged plane inside the sphere is opaque and white, outside the sphere it is opaque and green, and the projected image is translucent white. 138

7.3 The principle of the inverse stereographic projection applied to the cone and cylinder. 138

7.4 The projection on a cone's surface, as in Figure 7.2. 139

7.5 The projection on a cylinder's surface, as in Figure 7.2. 140

7.6 Projection on a cone (top row) and cylinder (bottom), shown their projection origin at the same place as in their equivalent (Figure 5.15) spheres (left column), and set higher for a stretching effect (right column). 141

7.7 Terrain flattening before inverse stereographic projection to a sphere, seen from the inside. From left to right: no flattening, gradual flattening after the transition to half the original height, and less gradual and complete flattening. 141

7.8 Projection of a hilly city without flattening (top left), and fully flattened before projection (top right). The differences in elevation between the two are shown in a color spectrum (green: no difference; bottom left), and magnified (bottom right). . . . 142
7.9 Inverse stereographic projection of a city onto a sphere, with an elevated position of the viewer. Projected are building footprints and street networks, colored by function, as well as the city's river. Buildings outside the sphere are dimmed but otherwise unchanged. Streets are slightly elevated over their original position above the terrain. . . . 143
7.10 Inverse stereographic projection of a city onto a sphere, seen from a pedestrian perspective. City geometry remains unaffected outside the sphere; streets are colored by elevation, and the 3D shapes of buildings are projected transparently. . . . 144
7.11 External and internal views of the inverse stereographic projection with different *radii* on the sphere, bisected by the base surface at its equator in the left column. A short *focus view extent* is shown in top row, with a long one below. The *intersection ratio* is smallest in the top right, larger in the bottom right, and at its maximum (50%) in the left column. . . . 145
7.12 The four linked primary parameters of the inverse stereographic projection: radius (r), projection intersection height (h), intersection angle (α), focus view extent (v). . . . 146
7.13 Automatic intersection angle: the same radius within rows, the same focus view extent within columns (increasing to the right and down). . . . 147
7.14 Automatic radius: the same intersection angle within rows (increasing to the right), the same focus view extent within columns (decreasing down). . . . 148
7.15 The interpolation (middle) from the original scene (left) to its projection on the sphere (right). . . . 148
7.16 The projection surface's vertical axis is bound to the viewer and moves with the virtual camera horizontally. Ray casts intersecting the surface are projected down to the original terrain using the same stereographic projection method. . . . 149
7.17 Inverse stereographic projection of a city, showing the controller used for navigation with its compass indicator and casting a ray to the projection for navigation. . . . 151
7.18 The map exploration system as used for the informal evaluation shown at three different zoom levels; fully zoomed out the whole map can fit inside the focus view extent and is thus not projected. . . . 152
7.19 Four deformation methods: Inverse stereographic projections on a cone (top left) and a sphere (top right), and bending along one axis (bottom left) or along a circle (bottom right) that is located at the cone's or sphere's intersection. The cone and sphere are not "equivalent" here, but adjusted for a more comparable inside view. . . . 154

7.20 The double stereographic projection described in section 5.3.7, shown with two different sizes of the projection sphere (top vs. bottom), and seen from inside looking toward the horizon (middle) and toward the zenith (right). The projection starts just before the true horizon, with a small overlap visible. . 155
7.21 Inverse stereographic projection of only the streets, rivers, and landmarks, seen from a pedestrian view. The cyan streets have been slightly vertically displaced relative to the magenta ones to show overlap. 159
7.22 A similar projection as in Figure 7.21, with a larger sphere and seen from an elevated position. This configuration showcases the multilayer capability of the system. 161
7.23 The current state of regular, "aerial"/overview (top) and AR (bottom) navigation assistance in consumer products. From Apple's June 2021 WWDC Keynote [App21]. 163
7.24 The double stereographic projection (section 5.3.7, Figure 5.17) applied to a globe of the Earth. The intersection occurs exactly at the horizon: the projection shows everything behind it, conformally distorted. 165
7.25 The double stereographic projection (section 5.3.7, Figure 5.17) applied to the Moon. The top left shows the intermediary step of projecting the original surface first to an intersecting plane, before projecting it again to the viewing surface. The other images show different configurations, including intersections happen before the true horizon. 166
7.26 A view of the Earth rising behind the Moon, overlooking Mare Smythii. Taken during the Apollo 10 mission on May 23, 1969. NASA index AS10-27-3896. . 167
7.27 The "Blue Marble" photograph of the Earth, taken by the crew of Apollo 17 five hours into their flight to the Moon on December 7, 1972, at about a distance of 29,000 km. NASA index AS17-148-22727. 169
7.28 Jackson Pollock in his studio in 1950 in a photograph taken by Hans Namuth. 172

9.1 An excerpt of the gaze data recorded during the study in section 6.2, showing which map layers participant №24's looked at in fine detail for each scenario. 182
9.2 An application of the inverse stereographic projection to a vertically-oriented half-cylinder, with the projection origin not being a single point above the viewer, but rather the infinite line at the top of the cylinder. Lines perpendicular to the viewing direction remain straight in their projection. 184
9.3 The MCV stack adapted to act as a overview+detail visualization over a 3D city model. Shown on the left as a model floating at table-height, and on the right as lying on the (virtual) floor. Blue lines from the corner of the lowest layer show from where in the model the layers are projected from. 185

10.1 Proximity: The matrices of dots are perceived as groups of rows or columns, just as the clumps of dots is perceived as consisting of tho groups because of spatial proximity or concentration. 189

10.2 Similarity: The matrix elements are perceived as rows or columns because of similar colors, shapes or orientations. 189

10.3 Connectedness: Elements that are close together or similar are not perceived as a group as much as those that are connected. 189

10.4 Continuity: Intersecting lines are perceived to continue their smooth shape and are not assumed to be taking sharp corners or be divided. 190

10.5 Symmetry: The same identical shapes appear more closely grouped when being on an axis of symmetry. 190

10.6 Closure: The red shape appears as a circle occluded by the rectangle and not as a broken ring. 190

10.7 Relative Size: The smaller blue areas are more likely to be perceived as one object. 190

10.8 Polar plots showing the density probability distributions of absolute and relative saccades in the viewport—the eye movement component—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**). 193

10.9 Polar plots showing the density probability distributions of absolute and relative saccades of the head—the HMD movement component—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**). 194

10.10 Polar plots showing the density probability distributions of absolute and relative total saccades—the gaze, as projected on a sphere—in degrees, represented as polar plots, for each MCV system (blit/grid/stack) and subgrouped by the participants' preference for an MCV system regarding its legibility (**top**), the frequency with which they play video games (**middle**), and the frequency with which they consult city maps (**bottom**). 195

10.11 (**A**bove: Isometric views from the side of a flat configuration of three layers, a stack with each layer perpendicular to the viewline (2), and a stack with each layer tilted at 40 degrees (3). Colored lines between illustrate differences in distance between camera and equidistant points on the map views: green is the average distance, a shift to red signifies less distance, a shift to blue more. (**B**)elow: the corresponding camera views from (**A**), with equal fields of view (tilted up in (**B1**)). All layers equal in size. 196

10.12 (**A**)bove: Isometric views from the side of a flat configuration of six layers (1), one stack with each layer tilted at 25 degrees (2), two stacks with each layer tilted at 35 degrees (3), and two stacks as in (3) but rotated to face the viewer (4). Colored lines illustrate distances as in 10.11. (**B**)elow: the corresponding camera views from (**A**), with equal fields of view (cropped sides in (**B1** and (**B2**)). All layers equal in size. 197
10.13 The shader graph implemented for the inverse stereographic projection on a sphere, cone, and cylinder, with all projection parameters listed in section 7.1.2 and an implementation of the geometry bending approach from [Lor+08]. 198

List of Tables

6.1 Effects of the system used, the scenario presented, and their interactions on time spent gazing at different data layers and in total, and on the number of times one of the data layers was mentioned (Mnt.) in explaining a choice (p-values). 124

7.1 Which high-level projection parameters are accessed by which axis of the touchpad under which condition, i. e. which *other* parameter is being automated (header row). 147



Colophon

This thesis was typeset with LaTeX 2$_\varepsilon$. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at http://cleanthesis.der-ric.de/.

Declaration

I hereby declare that I have completed this work solely and only with the help of the references mentioned in the bibliography.

Nantes, June 15, 2021

Maxim Spur

CENTRALE NANTES

Titre : Visualisation immersive de données géospatiales urbaines multicouches pour l'analytique visuelle et l'exploration

Mots clés : Données urbaines, géovisualisation, réalité virtuelle, analytique immersive

Résumé : Les données géospatiales urbaines englobent une multitude de couches thématiques et s'étendent sur des échelles géométriques allant d'éléments architecturaux aux réseaux de transport interrégionaux. Cette thèse étudie comment les environnements immersifs peuvent être utilisés afin d'aider à visualiser efficacement ces données multicouches de manière simultanée et à diverses échelles. Pour cela, deux prototypes logiciels ont été développés afin de mettre en œuvre deux nouvelles méthodes de visualisation de données, à savoir les « vues multiples coordonnées » et le « focus+contexte ». Ces logiciels tirent pleinement parti des possibilités offertes par le matériel moderne de réalité virtuelle tout en étant également adaptés à la réalité augmentée. Parmi les deux nouvelles méthodes présentées ici, l'une — une disposition verticale optimisée des couches cartographiques — a été évaluée de manière formelle dans le cadre d'une étude contrôlée auprès des utilisateurs, et l'autre — une approche de projection géométrique pour créer des vues panoramiques de type focus+contexte — de manière plus informelle à partir des commentaires d'experts du domaine qui l'ont testée. Si les deux méthodes ont montré des résultats prometteurs, l'étude formelle a plus particulièrement permis de comprendre comment les particularités des utilisateurs peuvent influencer la perception de la facilité d'utilisation de ces systèmes de visualisation ainsi que leurs performances.

Title: Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration

Keywords: Urban Data, Geovisualization, Virtual Reality, Immersive Analytics

Abstract: Geospatial urban data encompasses a plethora of thematic layers, and spans geometric scales reaching from individual architectural elements to inter-regional transportation networks. This thesis examines how immersive environments can be utilized to effectively aid in visualizing this multilayered data simultaneously at various scales. For this, two distinct software prototypes were developed to implement the concepts of *multiple coordinated views* and *focus+context*, specifically taking full advantage of the affordances granted by modern virtual reality hardware, while also being suitable for augmented reality. Of the two novel methods introduced here, one — an optimized, vertical arrangement of map layers — was formally evaluated in a controlled user study, and the other — a geometric projection approach to create panoramic focus+context views — informally through feedback from domain experts who tested it. Both showed promising results, and especially the formal study yielded valuable insights into how user characteristics can influence the perceived usability of such visualization systems and their performance.