Universität Rostock

Traditio et Innovatio

Aus der Professur für Geodäsie und Geoinformatik
der Agrar- und Umweltwissenschaftlichen Fakultät

Embodied Geosensification—Models, Taxonomies and Applications for Engaging the Body in Immersive Analytics of Geospatial Data

Dissertation

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

an der Agrar- und Umweltwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von M.Sc. Markus Berger

Rostock

Rostock, den 06.03.2023

Diese Arbeit ist gleichzeitig veröffentlicht in:

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK), Reihe C, Dissertationen, Heft Nr. 908, München 2023, ISBN 978-3-7696-5320-5, ISSN 0065-5325, www.dgk.badw.de

https://doi.org/10.18453/rosdok_id00004389

Prüfungskommission

Vorsitzender:	Prof. Dr.-Ing. Henning Bombeck (Universität Rostock)
Gutachter:	Sen.-Prof. Dr.-Ing. Ralf Bill (Universität Rostock)
	Prof. Dr.-Ing. Jörn Plönnigs (Universität Rostock)
	Prof. Dr.-Ing. Christian Tominski (Universität Rostock)
	Prof. Dr.-Ing. Florian Hruby (Berliner Hochschule für Technik)
Tag der Verteidigung:	30.06.2023

Danksagung

An dieser Stelle möchte ich allen danken, die mich während der Anfertigung dieser Arbeit unterstützt haben. Mein besonderer Dank gilt Professor Ralf Bill für seine langjährige und unermüdliche Unterstützung während meines Studiums, meiner Arbeit an der Professur für Geodäsie und Geoinformatik und darüber hinaus. Ihre Ratschläge, Ermunterungen zur Teilnahme an interdisziplinären wissenschaftlichen Konferenzen und konstruktiven Kritiken haben mir geholfen, mich akademisch und persönlich weiterzuentwickeln. Ohne Ihre Expertise und Ihre Unterstützung wäre diese Arbeit so nicht möglich gewesen.

Ebenfalls möchte ich mich bei Professor Jörn Plönnigs für seine großzügige Unterstützung in den finalen Phasen der Arbeit bedanken. Du hast mir vor allem in den entscheidenden letzten Wochen den Rücken freigehalten und damit sehr zum Erfolg der Arbeit beigetragen.

Nicht zuletzt möchte ich auch meiner Familie und meinen Freunden danken, die mir während der gesamten Zeit mit Rat und Tat zur Seite gestanden haben. Eure moralische Unterstützung, euer Verständnis und eure aufmunternden Worte haben mir viel bedeutet und mir auch in stressigen Zeiten als Anker gedient.

Abstract

Interactive maps are a constant companion in today's world, aiding us in tasks from wayfinding to making sense of complex ecological or demographic processes. They appear in everything from location-based apps, to highly specialized, dedicated data analytics tools. Meanwhile, the representation of these maps usually remains similar between these contexts—they are rendered in 2D or simple 3D views, and display their information visually.

As effective as these maps are, they do not fully exploit the multimodal way in which humans naturally examine spatial problems and make sense of their environment. Our reasoning is not only based on visual information, but on our capacity to explore space through movements, through action and reaction, through touch, hearing, smell, and taste. Modern spatially tracked and immersive computing systems like augmented and virtual reality headsets now allow us to experiment with applying such methods of reasoning to geospatial problems, by moving cartographic representations into the immediate spatial reach of our body.

This doctoral thesis explores the different ways in which such integrations of multisensory representation, geospatial data and body-driven interactions can aid us in analytical sense-making about real-world processes and phenomena. It starts with a broad investigation of an interdisciplinary array of research topics and synthesizes their aspects into a collection of taxonomies that can describe the most important design decisions that are involved in creating these "Embodied Geosensifications". The taxonomies are then integrated into a model and diagram language that can be used for high-level systems specification. This model is further applied to several practical examples from existing research, demonstrating that it can accurately specify different kinds of systems and aid in identifying design issues and their possible solutions. The presented research results are expected to promote more integrated and coherent thinking about such embodied geosensifications in the future.

Zusammenfassung

Interaktive Karten sind ein ständiger Begleiter in der heutigen Welt. Sie helfen uns bei Aufgaben von Wegfindung bis hin zum Analysieren komplexer ökologischer oder demografischer Prozesse. Sie finden sich in den verschiedensten digitalen Werkzeugen, von mobilen Kartenanwendungen bis hin zu hochspezialisierten Datenanalysetools. Die Darstellung dieser Karten ist in all diesen Kontexten in der Regel ähnlich—sie werden in 2D- oder einfachen 3D-Ansichten gerendert und bleiben dabei rein visuell.

So effektiv diese Karten auch sind, nutzen sie die multimodalen Arten in denen Menschen räumliche Probleme verstehen und ihre Umgebung wahrnehmen nicht vollständig aus. Unser Denken basiert nicht nur auf visuellen Informationen, sondern auch auf unserer Fähigkeit den Raum durch Bewegungen, durch Aktion und Reaktion, durch Berührung, Hören, Riechen und Schmecken zu erkunden. Moderne immersive Technologien wie z.B. Augmented- und Virtual-Reality-Headsets ermöglichen es uns nun, solche Denkmethoden auf geografische Fragestellungen anzuwenden, indem wir kartografische Darstellungen in die unmittelbare Reichweite unseres Körpers bewegen.

Diese Dissertation untersucht die verschiedenen Möglichkeiten, wie eine solche Integration von multisensorischen Darstellungen, Geodaten und körpergesteuerten Interaktionen uns bei der Analyse von Prozessen und Phänomenen der realen Welt unterstützen kann. Sie beginnt mit einer breit angelegten Recherche eines interdisziplinären Spektrums von Forschungsthemen und fasst deren Aspekte in einer Sammlung von Taxonomien zusammen, welche die wichtigsten Designentscheidungen bei der Erstellung solcher "Embodied Geosensifications" klassifizieren können. Die Taxonomien werden dann in ein Modell und eine Diagrammsprache integriert, die für die Spezifikation von Systemen auf einer hohen Abstraktionsebene verwendet werden können. Dieses Modell wird auf mehrere praktische Beispiele aus der bestehenden Forschung angewandt, um zu zeigen, dass es verschiedene Arten von Systemen spezifizieren kann und schon im Entwurf dabei hilft Probleme zu identifizieren und mögliche Lösungen zu finden. Die vorgestellten Forschungsergebnisse sollen dabei dazu dienen, dass zukünftige Beispiele solcher Systeme ihre einzelnen Aspekte klarer und besser miteinander integrieren können.

Contents

List of Abbreviations

ADSR	**A**ttack, **D**ecay, **S**ustain, **R**elease
AR	**A**ugmented **R**eality
AV	**A**ugmented **V**irtuality
CNOSSOS-EU	**C**ommon **No**ise A**ss**essment Meth**o**d**s** in **Eu**rope
DEM	**D**igital **E**levation **M**odel
DR	**D**iminished **R**eality
GeoIVE	**Geo**visualization **I**mmersive **V**irtual **E**nvironments
GIS	**G**eographic **I**nformation **S**ystem
GLTF	**G**raphics **L**anguage **T**ransmission **F**ormat
HCI	**H**uman **C**omputer **I**nteraction
HRTF	**H**ead **R**elated **T**ransfer **F**unction
ISO	**I**nternational **S**tandards **O**rganization
ITD	**I**nteraural **T**ime **D**ifference
IID	**I**nteraural **I**ntensity **D**ifference
LBS	**L**ocation **B**ased **S**ervices
LED	**L**ight **E**mitting **D**iodes
LFO	**L**ow **F**requency **O**scillators
LOD	**L**evel **O**f **D**etail
MR	**M**ixed **R**eality
NeRF	**Ne**ural **R**adiance **F**ields
OGC	**O**pen **G**eospatial **C**onsortium
SLAM	**S**imultaneous **L**ocalization **a**nd **T**racking
STEM	**S**cience **T**echnology **E**ngineering **M**athematics
TENS	**T**ranscutaneous **E**lectrical **N**erve **S**timulation
TIN	**T**riangulated **I**rregular **N**etwork
UI	**U**ser **I**nterface
UML	**U**nified **M**odelling **L**anguage
VAC	**V**ergence **A**ccomodation **C**onflict
VE	**V**irtual **E**nvironment
VGE	**V**irtual **G**eographic **E**nvironment
VR	**V**irtual **R**eality
XR	**E**xtended **R**eality

Chapter 1

Introduction

Deliberating about our relation to the world around us has been a favourite pastime of philosophers, scientists and poets for centuries. Intellectual debates are still being fought over how deeply our own cognition is integrated into our immediate surroundings and what role the different parts of our body play in this integration. There are hundreds of theories, many of them mutually exclusive, about how we conceive, memorize and understand spatial context. It remains unclear if it is even possible to be a thinking, conscious being without having "a body" within "an environment".

One of the most important—and simultaneously one of the oldest—steps in expanding our spatial context is the creation of maps. Cartography takes what is too far for our sensory faculties to recognize and moves it closer to us. This basic fact has not changed from ancient cave drawings to modern virtual 3D globes. However possible it may be to understand oneself without spatial perception, we can never understand "the world" without bringing it closer to our own bodies. The way we draw our maps has deep influences on our understanding of the environment, other people, geopolitics and many areas of modern life. Cartography, in a way, can be seen as visual communication about the world that is beyond our immediate perception. It cannot be overstated how important that makes cartography for our understanding of the planet and the global systems we operate in.

This makes it all the more surprising that most of our modes of utilizing computational power to aid this understanding still end up mirroring flat 2D maps printed on paper. Even when rendering 3D maps, they are often constrained to being 2D surfaces with height-based information added on top. Why is spatial information not something we regularly put into the world around us, something to feel, hear and manipulate as we wish, even though technology would allow us to? What makes flat maps on flat screens such a powerful concept? Apart from just familiarity, there are multiple commonly cited reasons for this (lack of) development:

1. The visual sense is very effective for taking in large amounts of information.
2. For many types of data and analysis, dimensional reduction into two dimensions is efficient.
3. 2D displays are simpler to build than 3D displays.
4. 2D input devices are simpler to build than 3D input devices.

In a world where 2D maps are sufficient, and 3D display and interaction technologies are difficult, we have all grown accustomed to the former. But simultaneously, we have always dreamed of going further, even before we developed display technologies that allow us to perceive beyond a flat plane. Conventional maps make the environment accessible, but they also disembody it, making us unable to directly intervene without employing abstract interfaces like mouse and keyboard. Why should maps not be something we can touch, move and shape? Experimental haptic and force-feedback controllers regularly try to capture new markets for human-computer-interaction, such as training and recreation. Why should maps not be something to listen to? Modern sound interfaces are highly advanced in displaying spatialized audio information. There is an appetite for new additions to the canon of cartographic knowledge—what needs to be figured out is how to use these new technologies in the geospatial, cartographic context.

1.1 Motivation and Research Questions

Whenever interest arises in employing non-standard technologies like immersive interfaces or non-visual displays for geospatial data analytics, those that are supposed to design these systems quickly encounter an issue not often found in geovisualization today: a lack of precedent. Often, there will be a set of data and a goal of what kind of information should be extracted from it. In conventional cartography, we can fall back to a large number of existing techniques and traditions. The design idea is quickly drafted by invoking and comparing commonly understood reference solutions to each other. But once we want to make use of new technologies from domains like Virtual Reality (VR) or new non-visual display systems, such techniques and traditions are often missing entirely. Without these building blocks, it can be exceedingly difficult to identify possible roadblocks as well as opportunities before already being deep into the implementation of a system—at which point it is often too late to naturally integrate them.

Lets consider a collection of point values over a city, representing measurements. The pathway to a 2D geovisualization here is simple: Find a 2D base map. Convert the data into an appropriate GIS (Geographic Information System) format that works with the chosen visualization software. Superimpose the data by picking a fitting color palette from a predefined list. Map the measurement values onto this palette. Then consider more advanced cartographic questions like label placement, following well-established guidelines. Once the map is done, enable a set of predefined interactions, such as retrieving numerical data values by clicking on a point.

Now we imagine the same situation, but the data is supposed to be shown while the user is located in the 3D model of the city. Maybe the data display is an ancillary part of a larger system that lets users explore virtual cities. Maybe the data needs to be shown in a real-life context, through augmented reality (AR) technology. Suddenly, there is no precedent. Should the points be shown on the ground below the user or floating above them? Should the points be transparent so the world behind them is still visible? What happens when the user enters a densely built area and most points are completely obscured? How do we access values that are far away? How do we select values or interact with the data when all our VR motion controller offers is one button and a virtual laser pointer? What if the very limited screen real estate is already used up and the data needs to be shown through purely non-visual means?

Currently, there are only scattered experiments and disparate models of how to approach these problems. Even with a solid technical foundation in immersive display technologies, applying them to data visualization problems is challenging. Some of the issues that lead to the conception of this thesis stem from exactly such challenges. Berger and Bill (2019) describes the implementation of a system that tries to show what the values visualized on an urban traffic noise map really mean. It achieves this by placing the user into a virtual city model and playing the estimated traffic noise back as an audible sound. But there was a problem: while the audio playback could be developed according to mostly existing concepts from the field of sonification, the noise map itself still needed to be visible to give a broader spatial context. But the ground-level viewpoint of an immersive virtual city introduces multiple forms of occlusion, such that data gets less visible the further it is away. With no true solutions to this issue at hand, subsequent research investigated multiple ways of tackling this problem, from a sound and haptic-based display for distant values, to interaction-based methods. During the sound-based investigation in Berger (2020), it became clear that we were missing a clear conception of sound-based data display in immersive environments—and that this extends to all the other primary senses too. While visualization has many appropriate taxonomies for such axes, the broader category of data display that spans all the human senses, in this thesis called *Sensification*, still lacks them. During the interaction investigation in Berger (2021a), the same issues became clear for the concept of *Embodied Interactions*—those user interactions that use movements, gestures and tangible objects for input. Similar problems were also encountered concerning the quality of geospatial data necessary for display in immersive systems.

However, despite these challenges, there has been a clear and active demand for embodied and multisensory systems within the literature and on research agendas for many decades, even before the current hype around immersive virtual reality displays. This is usually done under the larger banner of *Immersive Analytics* (Ens et al., 2022), but sometimes more specifically for the geospatial case (Slocum et al., 2001). Ultimately, the hypothesis driving this field is as such: 2D maps and visualizations are powerful for what they do, but immersive analytics systems can solve entirely new problems or allow us to tackle old problems in new ways.

So far, the issues of multisensority, immersive geovisualization, and embodiment are often separate bullet points on such research agendas and only occasionally intersect. This thesis will be an attempt to explore this overlap, from here on called *Embodied Geosensifications*. The underlying hypothesis is that this combination has the potential to be a potent tool, but our lack of clear conceptions of how these aspects can integrate with each other often obscures what design pathways can even be taken. To develop a model of these possibilities, we identify four main research questions. The first three questions concern the individual aspects considered on their own: human *senses* (multisensority), immersive geospatial *data* representation (geovisualization) and body-driven *interactions* (embodiment). The last question concerns the conceptual combination of these aspects:

1. **Senses**: What are the primary sensory modalities of the human body that we can display data to? Which modality should be targeted for what data?
2. **Data**: Do we have to evolve our conception of geospatial data for immersive systems? What are key factors in this evolution?
3. **Interactions**: What kinds of interaction can there be between the human body and geospatial data?
4. **Model**: How can we model embodied geosensifications in a practical way?

1.2 Structure and Content Overview

This thesis will start with a survey of the current state of the art of (geo-)spatial information visualization in Chapter 2. This chapter covers all topics that are relevant for the creation of taxonomies and models that actually take into account the current state of technology—what are the current frontiers in 2D data visualization and cartography, and how far have we gone with 3D, multi-sensory display technologies past and present? What solutions are examples of truly spatial and sensory data sensification and how well do they seem to work?

Equipped with that knowledge, we will follow the four research questions. The first three questions make up Chapter 3, which attempts to clearly define a set of possibilities, as well as the technological, cognitive, and physiological considerations for each of the three aspects.

In Chapter 4, we will attempt to rework the defined taxonomies into a complete model of multi-sensory, immersive, embodied geospatial analytics systems. As such, this chapter has to provide the glue between the previously established concepts. The individual steps frequently refer back to the concepts of senses, data and interactivity.

In order to show how the developed model can be applied, Chapter 5 will take a more practical turn and return to the prototypes that lead to the development of this thesis. These examples will be relayed to the concepts that their creation helped influence or that were important in their creation. They will also be retroactively subjected to the model developed during the earlier chapters, in order to create a varied collection of examples that cover most aspects of the developed model and diagram, and to demonstrate that the diagram is expressive enough to handle very different kinds of systems.

For a visual overview of how the ideas in this thesis develop from chapter to chapter and section to section, see Figure 1.1.

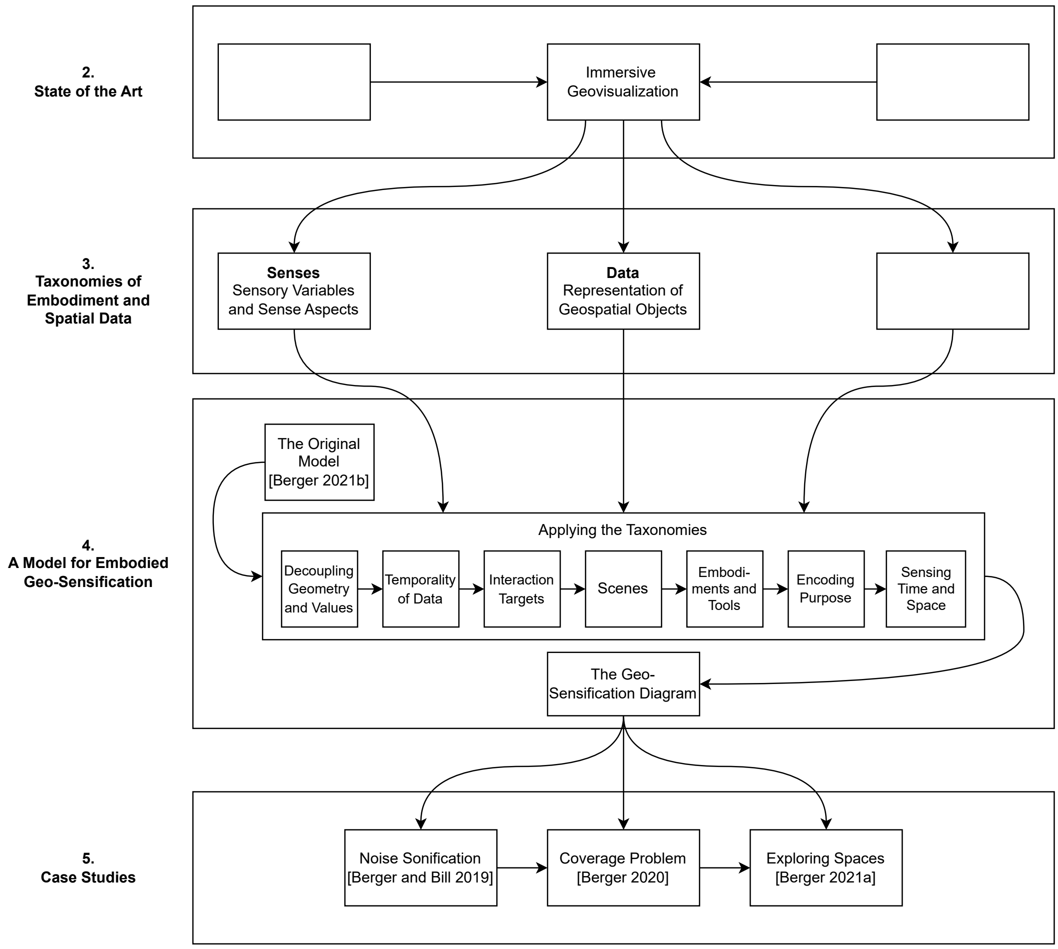


FIGURE 1.1: An overview of the content of this thesis. The four main boxes represent Chapters 2 through 5. The included boxes show the sections of each chapter. The arrows represent how the sections iteratively build upon the concepts established within them.

Chapter 2

State of the Art

2.1 Visualization and Visual Analytics

With visual displays being the dominant form of data presentation, this chapter will start at the most common point of entry: Visualization.

The term visualization is used in many different contexts and can refer to many applications. In a way, every form of graphics displayed on a computer screen or any drawing or print visualizes something. However, when we speak of visualization in a scientific context, we are usually referring to the concepts of *information visualization* or *scientific visualization*. The former tends to refer to the visual display of abstract data set (economic data being perhaps the most prominent example), the latter to visual display of data about some physical fact or process.

Whatever kind of fact is visualized, there is a set of criteria most visualizations try to follow: (Card, 2009) (Schumann and Müller, 2013)

Expressiveness A visualization should only display what is actually part of the relevant data set.

Effectiveness A visualization should only display what the human visual system is capable of perceiving and can perceive well.

Appropriateness A visualization should only be made if the effort to create it does not vastly exceed its usefulness.

These criteria are of course very generalized—attaining a visualization that fulfills these guidelines is usually a process that requires far more specific workflows. The main way we think about visualization today is the visualization pipeline as first proposed by Haber and McNabb (1990), shown in Figure 2.1.

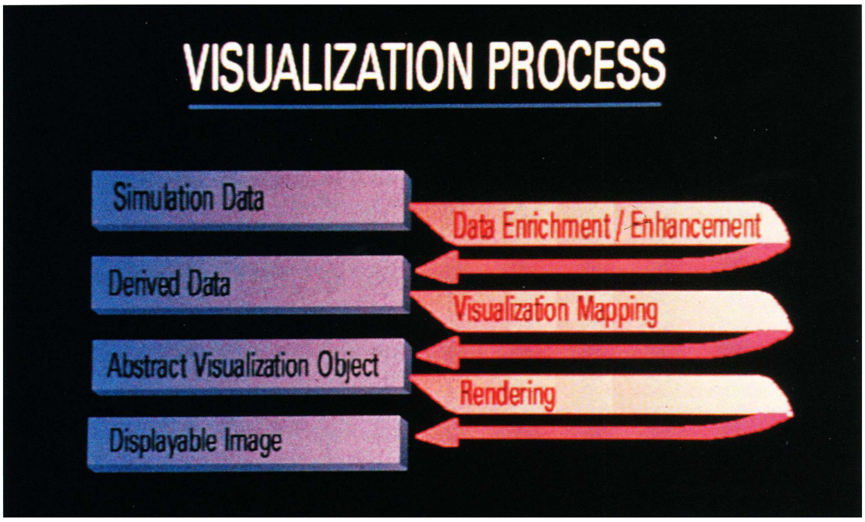


FIGURE 2.1: The visualization pipeline as it appeared in Haber and McNabb (1990).

It has been expanded multiple times and applied to multiple sub-disciplines of visualization, however the general concept stays the same: We take a set of data that we have enriched in some way (for example by data analysis methods), we prepare that enhanced set of data for the visualization by filtering or reducing it, then we map it to a representation that can usefully be displayed on a screen, and then we render that representation. A more modern pipeline that explicitly includes all these steps can be found in Dos Santos and Brodlie (2004) and is shown in Figure 2.2.

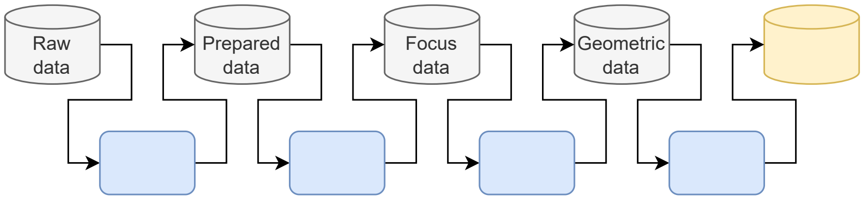


FIGURE 2.2: The visualization pipeline as it appeared in Dos Santos and Brodlie (2004). (adapted)

One fact of note in both of these pipeline models is that they are shown to be unidirectional. A series of processes is completed, which then creates one result. There was a time when visualization worked in this way, but this can hardly be said to be the case today. Most visualizations are *interactive* in some way, to the point that definitions of the term visualization often already specifically include the concept of interactivity. Thus, what is called a pipeline could now perhaps be more properly called a *visualization loop*, or as Dwyer et al. (2020) put it, a "sense-making loop". Especially the steps of filtering and mapping are iterative and contain feedback. Sometimes this loop is even split towards the end, for example to create multiple renderings with different filters and mappings applied on the same set of data.

Decades of research work and visualization practice have led to a foundation of interaction techniques that are basically usable off-the-shelf today. Examples for such techniques are:

- Dynamic Query (Shneiderman, 1994). Users can adjust the filter that that is operating on the data set interactively and in real-time.
- Focus and Context (Bjork and Redstrom, 2000). Users can focus on a specific part of the data, while the broader context still remains visible.
- Coordinated Multiple Views (Roberts, 2007). Multiple visualizations are shown simultaneously and react in unison when an interaction happens on one of them.

One can find these techniques in many modern business dashboards and data visualization programming libraries. They can be seen as the building blocks that make visualizations interactive. Another kind of visualization canon are the common visualization types we almost reflexively apply any time we face a data analysis problem. These are usually called charts or graphs and denote a specific tradition of mapping data to visual artefacts. Examples include bar charts, line graphs, scatterplots and more. A visualization type of special interest for this thesis is the *map*, which specifically encodes geospatial data.

If we extend the filtering or data analysis steps of the pipeline to be interactive processes that are integrated in the same interface as the rendered data, we move into the highly interdisciplinary territory of *Visual Analytics* (Keim et al., 2006). Here, we integrate visualization with dynamic use of more classical data analysis steps (statistical analysis and other computational methods) and human computer interaction (HCI), to create intuitive, effective computer interfaces in which we can take a set of raw data and apply all steps of the visualization pipeline dynamically to probe and interrogate it. This process aims to attain knowledge about a given set of data more effectively than through the use of just statistical analysis or visualization.

After establishing some foundations, we now move on to the current state of the field. Cui (2019) offers a very thorough introduction to the evolution and contents of the term visual analytics, as well as of the challenges that the field faces. At the core of their conception of modern visual analytics is the concept of dimensionality and dimensionality reductions. This includes basic questions like what to do with highly multidimensional data sets, or when 3D visualizations are appropriate in comparison to 2D visualizations. This is caused by the data sets in question growing larger and more complex as our methods of data collection improve.

There is however another solution to multidimensional data that does not necessarily involve a reduction of dimensions. As technological progress creates novel display and interaction technologies, new dimensions of display beyond the conventional 2D screen become available. This can ranges from large touch screen arrays to experimental interfaces like smell displays. We have to ask what is more effective—using transformation and reduction techniques on the data or adding more parallel display channels? There is no final answer to this question, because it is highly contextual. It is also the wrong question to ask at this point in time, because the field of "visualizing" through such immersive and multi-sensory means is still barely developed. A better question might be: What are they ways in which we can utilize the new display and interaction dimensions that current hardware unlocks for us?

To begin answering this question, we will first move deeper into one specific sub-discipline of visualization, to then extend this sub-discipline back out into the multisensory realm. Considering the geospatial focus of this thesis, this sub-discipline is of course the one concerned with geospatial visualizations in the form of *maps*.

2.1.1 Cartography

"*Viewed in its development through time, the map details the changing thought of the human race, and few works seem to be such an excellent indicator of culture and civilization.*"

Norman J.W. Thrower, Maps and Civilization (J. W. Thrower, 2007)

The study of cartography is an old one and far too complex to fully encompass in a thesis like this one. Certain specific concepts and topics in cartography will become relevant in Chapter 3 and will be discussed there. This section will only briefly look at some current developments to then move on to how cartography relates to modern visualization specifically.

There are multiple active research directions in the cartography community. For one, there are specific technical problems, like label placement (Rylov and Reimer, 2014) or rendering and styling (Meier et al., 2014; Dübel et al., 2017). Many of these problems can be considered solved for most common use cases, but still require challenging research in many corner cases. Another example would be map projections, which are well established in everyday use, but are still a current topic in planetary cartography, where researchers try to answer questions like how to design projections for astronomical bodies with non-standard shapes or axis of rotation (Pędzich, 2019).

Another direction in cartography is the analysis of existing maps, like for example historical maps. This involves the use of ontologies, computer vision algorithms and historical knowledge. A similar direction is the use of existing image material in the form of aerial and satellite imagery—how do maps transform if we do not fully control the symbology and thus every pixel, but superimpose additional information onto a real-world image? This has some overlap with the AR and VR research which will be discussed in the following sections. Here, we for example superimpose labels on building surfaces in the user's field of vision (Postert et al., 2021) or create virtual environments purely from 360 degree images and then extend them with textual labels as well as other multimedia data (Westphal and Berger, 2020).

There is also the direction of critical cartography and other sub-disciplines that deal with the sociological, material and historical consequences of cartography and how cartography is done. Here, we often find a deep relation to other fields like philosophy and critical geography. This research is not of immediate importance in this thesis, as we are mostly focused on technological aspects—but especially the topics of embodiment and conceptions of space (see Section 3.2) had their roots in these disciplines before they ever became relevant to engineering and computer science.

Perhaps the most relevant however are two other research directions in cartography: First, the semiotics of maps and how they interact with geographical data, usually manifested in the form of map symbologies in GIS software, as well as Jacques Bertin's visual cartographic variables (Bertin, 1967). These topics will be dealt with in detail in Section 3.1 of this thesis. These concepts have been established and refined for decades. Most of the evolution that is still present within them is happening at the intersection with the topics of the following sections—immersive technologies and other novel display systems.

Far more active is the second relevant research direction: Interaction design and user experience research. Here, the field of cartography directly intersects back with visualization and visual analytics, as it moves into the larger context of interactive, dynamic software systems, like the mapping applications we use everyday on our computers and phones. One large review study by Roth et al. (2017) sums this up neatly by starting with the following sentence: "The possibility of digital interactivity requires us to re-envision the map *reader* as the map *user*." They name data visualization and HCI directly as the main interdisciplinary influences on cartography, right next to geography and psychology. For years now, a large part of the cartographic discipline has been quantitative user study-driven research interrogating very specific questions of interaction design and cognition.

Because the questions that are being answered here are usually very context-specific, this kind of cartography has been shifting very quickly as new classes of consumer devices become available. New generations often engage very differently with maps. Where their parents might have learned to read maps out of pure necessity when travelling to unfamiliar locations by car, a child's first run-in with maps today may just be the "minimap" displayed in the corner of their favourite video game—which has proven implications for how maps are perceived (Horbiński and Zagata, 2022). Modern cartographic software has to take these trends into account. The same will be true for the kind of systems that are the focus of this thesis.

9 of the 18 questions in the ICA's 2017 paper on the future of cartographic research (Griffin et al., 2017) can be directly mapped into the purview of this specific research drive in the field of cartography. It is also important to note, that in this interactive usability-driven context, the term cartography has to some degree been left behind. Most of the literature, when referring to interactive cartographic applications, prefers the term "Geovisualization" (Çöltekin et al., 2017) or less commonly the term "Geovisual Analytics" (Andrienko et al., 2010).

2.1.2 Geovisualization

The term Geovisualization has been in use for longer than this current cognitive drive of research and has been used to refer to many digital forms of cartography. Its specific meaning as an interactive, data-driven and explorative form of digital cartography came about in the 1990s, spearheaded by Alan MacEachren MacEachren (1995), building on foundations laid by Bertin (1967) and others, and fueled by the rapid progress of scientific and information visualization research at the time.

Interestingly, these publications coincided with a massive boom in interest and research around virtual reality (see Section 2.2) which results in many sources from the time, like much of the ICA's agenda on cognitive and usability issues in geovisualization from 2001 (Slocum et al., 2001), being focused around the intersection of geospatial data and immersive virtual environments. The agenda laid out then remains useful to this day, as six major research themes were identified:

1. Geospatial virtual environments (GeoVEs)
2. Dynamic representations (including animated and interactive maps)
3. Metaphors and schemata in user interface design
4. Individual and group differences
5. Collaborative geovisualization
6. Evaluating the effectiveness of geovisualization methods

Especially virtual environments, including their effectiveness, how to collaborate in them, and their interface design are still areas of active research, some of them having only become relevant again after they had fallen out of fashion for years. This is well illustrated by an agenda paper from Çöltekin et al. (2017), which includes very similar categorizations of research drives, but is much more current. Meng (2020) goes as far as saying that "user experiences of geovisual analytic approaches are far from being systematically studied."

There has also been progress on the third point in Slocum et al.'s agenda, i.e. progress in metaphors and schemata (in short: guidelines) for user interface design. Roth (2013b) explicitly invokes Bertin's and MacEachren's visual variables as historically helpful visualization primitives (although he refers to them as *representation primitives*) and attempts to empirically discover similar primitives for interaction design in maps. This is done by structured interview with geovisualization design experts. It is interesting to note here that Bertin's original variables were not at all empirically derived, but were entirely based on Bertin's professional experiences. Empirical study of these variables as well as additions to the models were done post-hoc by other research groups (MacEachren, 1995).

In his empirical study, Roth is trying to answer questions he already posed earlier in Roth (2013a), in the form of a research agenda for "cartographic interaction science"—a term which is rarely used in the literature, but might be useful in decoupling the interactions from the visualization where necessary. The agenda ends up not looking too dissimilar from the agenda shown earlier, but there is one interesting observation baked into it, which will inform assumptions made in Chapters 3 and 4: Every interaction includes three components—the user, the visualization and a mediating computing device. In order to design a useful interaction, we need to keep all of these components in mind and constantly ask ourselves questions like "When do we show what visualization?", "How exactly do we interact at what moment?", "What is the specific use case?", and many more. After articulating some of these questions in Roth (2013b), he goes on to focus in on the specifics of "How?", i.e. what individual interaction exchanges can we offer users to support a useful interaction strategy? He considers multiple existing taxonomies of interaction by conceptualizing the process of interaction as the use of interaction *operators* on visualization *operands*, which is used to fulfill *objectives* in order to reach a *goal*, and then claims that any interaction taxonomy can be either based on the *operand* (visualization-based), the *operators* (interface-based) or the *objectives* (intention-based). Roth goes on to use his empirical study to reincorporate concepts from other taxonomies, both combining them and resolving some of the conflicts, thus reaching a synthesis that makes this specific taxonomy very useful for the rest of this thesis. The specifics of this taxonomy will be further discussed in Section 3.3 of this work.

Another issue not mentioned in the earlier agendas is the move from purely geospatial to spatiotemporal data. As advances in sensing and database technology have been implemented around the world, large multidimensional geospatial time series need to be visualized. Here, entirely new visualization techniques are still in active development. Some examples include methods visualizing trajectories (Tominski et al., 2012), flows (Boyandin et al., 2011) and events (Sheidin et al., 2017).

2.1.3 3D Geovisualization

With their roots firmly set in the research of virtual environments, geovisualizations are already often three-dimensional, even if much of the underlying GIS data tends to be stored as two-dimensional geometries. However, there are some trends that warrant direct notice.

First, one kind of spatial data that is reliably three-dimensional is large-scale simulation data, as found in meteorological and flood-risk simulations (Rydvanskiy and Hedley, 2020). Here, the full accuracy of both the simulation itself as well as of the visualization can only be attained by moving into the 3D space. Older two-dimensional methods are often pushed to their limits by modern requirements for simulation accuracy.

Secondly, there is one of the biggest research drives in all of geovisualization, often forgotten in agenda papers, as it is usually driven by governmental or corporate initiatives: large-scale 3D models, usually either scanned from aerial imagery or mass-collection of street-level images (Tan et al., 2020), or procedurally generated from existing geospatial data sets and large rule-based systems. Between projects like Google Maps and Esri CityEngine, mass adoption of CityGML models and more, there are huge implications for "in-place" geovisualization that can be dynamically utilized and embedded into these models, thus enabling the analysis of data in its true spatial context.

Perhaps one of the biggest hurdles for mass adoption of 3D geovisualization is the main issue with any form of 3D visualization: Occlusion and Distortion. The dimensional reduction to 2D often employed in cartography, as well as scientific and information visualization, is to a large degree caused by these two problems. It is often better to simplify the view on the data, than to require the user to deal with an interactive virtual camera in order to get a complete view of the data. For a number of years, it was business data visualization that gained some infamy for committing similar mistakes through the excessive use of 3D-features like shadows, depth and oblique viewpoints. We need to be careful that the reflexive use of 3D tools in our mapping applications does not suffer the same fate—there is now a plethora of research into dealing directly with exactly such problems (Deng et al., 2015; Zhang et al., 2016; Röhlig et al., 2017).

Most of the advantages and issues described in this section are further amplified once we also move our display and interaction devices into 3D. When we are wearing a VR headset and holding handheld motion controllers, suddenly not only the environment and the data itself can occlude our view on the data, but also our virtual body and the 3D user interface (UI) elements. In the following sections, we will take a look at the basic concepts and technologies present in this space and then reincorporate them with what we have already learned about geovisualization.

2.2 Spatial and Immersive Computing Technologies

Whenever a new technology accumulates enough hype to exit the realm of experimental research prototypes, it will quickly find itself in a constant state of terminological confusion, trapped between scientific definitions, corporate marketing and common parlance. And while most seemingly settle on some basic terms after some years of popularity, the technologies here have experienced no such equilibrium. Perhaps the most settled starting point are the terms of Virtual Reality (VR) and Augmented Reality (AR). In its most basic form, VR is about entering a different, virtual world while shutting out the real world. AR on the other hand is about adding digital objects to the real world.

This basic distinction is still very congruent with the originally established scientific definition by Milgram and Kishino (1994). Here, the authors propose a "Mixed Reality Continuum" with the real world on one side, and virtual reality on the other. This is a continuum of *immersion*, i.e. how deeply our sensory information is immersed in either the real world (i.e. real world stimuli reach the sensory organs as usual) or the virtual (i.e. real world stimuli are replaced by synthetic stimuli). In this model, AR broadly refers to the whole left side of the continuum—sensory states in which most stimuli originate from the real environment, but some are replaced on the way. Ultimately every stimulus is of course "real", the replacement here refers to some form of display device that blocks out anything coming in from an "outside" (which is contextual depending on the hardware used) and instead induces different information into the sensory organ it displays to. Milgram also proposes to call the right side of the continuum "Augmented Virtuality" (AV), i.e. a state where most but not all sensory input is replaced, though the term has not caught on in either science or consumer spaces.

The problem is that this proposed state of full immersion that Milgram calls Virtual Reality, is not actually what people mean when referring to VR today. Instead, what they refer to are usually head mounted displays (HMDs) that block the wearers complete field of view and then display a view into some virtual environment, usually combined with headphones and handheld controllers. Both HMD and handheld controllers are usually tracked in 3D space to offer the user an illusion of moving through a stable virtual world that tangibly exists in reference to their body. A full sensory immersion display like the one proposed by Sutherland et al. (1965), in which "A chair [...] would be good enough to sit in", is not what Virtual Reality is or even can be today. Modern VR is thus both less than what Paul Milgram defined, but also much more. Many today would not even consider a HMD that can display a virtual world enough to constitute VR, as the handheld controllers or body tracking done by modern systems has become so ingrained in the term and how it is viewed.

To make matters worse, the term Mixed Reality originally coined by Milgram and Kishino (1994), which used to be the headline for the whole field, is used outside of clear scientific contexts as little more than a synonym for versions of AR or AV in which the tracking and display is of especially high fidelity, usually involving head mounted displays (HMDs) as opposed to smartphones or simpler display glasses. Into this now vacant space of a good name for the whole technological space rushed other terms such as "Extended Reality" (XR) (Çöltekin et al., 2020) or "Spatial Computing" (Greenwold, 2003). Extended reality however seems like a synonym for AR, but one that would exclude certain related technologies such as "Diminished Reality" (DR) (Mori et al., 2017), in which real stimuli are replaced in such a way that it feels to the viewer as though they are still seeing the real world, but an actually existing object or stimulus is missing.

Ultimately, all of these terms convey different parts of this larger space of systems that combine spatially tracked, usually handheld or worn user interfaces with spatially tracked, depth-enabled displays in order to believably place interactive virtual objects into the users perception. Because of the spatial and multisensory focus of this work, we have to be much more specific about how the senses are targeted by displays than most audiovisual VR applications. Even papers that try to offer a broad overview of the field often specifically exclude senses other than sight or hearing, as the display and interaction technology serving them is still quite experimental (Çöltekin et al., 2020).

In addition to the senses, we have also have to consider our focus on geospatial data. Geospatial virtual worlds are unlike other virtual worlds in that they have a clear relation to the real world and thus not only track the user locally, but automatically place them into some form of geospatial reference. Considering the multiple stacked conceptions of spatiality in such applications, this thesis will refer to the breadth of technologies in this field as "Spatial Computing Technologies" and only use terms like AR, VR, XR or others when a paper or application is built with these specific distinctions in mind, usually when referring to visual display technologies. Spatial Computing neatly encompasses spatially tracked display devices, spatially tracked interaction devices, location-based services (LBS), the user's body moving through space, the different levels of tracking technology as well as the 3D interfaces popular within the applications. Because of the overlap with terms from the world of GIS and geovisualization, such as spatial analytics, we will always explicitly utilize the "Geo-" prefix when referring to such methods. An interface or display device is *spatial* as soon as it has some sort of local reference or tracking. It is *geo*-spatial once a global reference is present and relevant.

2.2.1 Display Technologies

As one of the main focuses of this thesis is the use of multiple senses for data display, we need to be aware of the basic facts of displaying information to each of the senses. While some of the information will be given once it becomes important in Section 3.1, what follows will be a general introduction into the state of the art for each of the five primary senses.

2.2.1.1 Visual Displays

Everyone reading this work will be familiar with different kinds of visual displays, especially of the 2D variety. Most commonly available displays at the time of writing are based on different types of light-emitting diodes (LEDs) whose light either acts as a backlight for liquid crystal panels or directly displays color values. Older technologies sometimes use moving electron beams that are fired onto a fluorescent display layer, thus creating either a pixel raster if the beams are regularly pulsed or vector graphics when the beam is continuous.

The most important aspect of visual displays for spatial computing is how they can be made to enable depth perception in virtual environments. There are multiple technologies in use to enable stereo-viewing, including systems that do this on one display, either through autostereoscopy or signal multiplexing combined with special stereo glasses that separate two slightly different images in such a way that each eye only sees one of them. In current spatial computing, the most common visual displays are HMDs, in which there is usually one high-resolution display for each eye, aided by a system of lenses to enable the user to focus and enable distance perception. The current static lens systems in these HMDs are not perfect however and can cause strain through effects like the Vergence-Accomodation Conflict (VAC), which can exacerbate existing difficulties with VR technology, like the motion sickness often caused by tracking latency issues (Yang and Sheedy, 2011).

A still more experimental class of visual HMD displays are the see-through displays used for AR applications. Here, either video cameras or optical systems are used so that the user can see both the real world as well as virtual worlds. If these worlds move in unison (through a technical process called registration), these displays can induce a sense of virtual objects existing fully within the real world (Azuma, 1997).

Even with the remaining issues in HMDs, visual displays are broadly our most advanced displays. The other senses, with the exception of the sense of hearing, are usually far less commonly "displayed to", because the technology to do so is simply not developed enough. A large list of currently available displays ordered from non-immersive over semi-immersive to fully immersive can be found in Çöltekin et al. (2020).

2.2.1.2 Audio Displays

Audio displays are as ubiquitous as visual displays and often naturally accompany them. Like visual displays, they can either be blocking or be "hear-through". Hear-through displays are much easier to design than see-through displays, as audio displays can be constructed in a way that blocks very little or none of the incoming sound, as is the case with loudspeakers and open-backed headphones, or simply record and replay the sound as needed, as is the case with many hearing aids or modern earbuds and close-backed headphones. Blocking audio displays either do so passively by putting an appropriate physical barrier between the ear canal and the outside environment, or they can be noise-cancelling, in which case noise is detected and then immediately eliminated by playing noise with an inverted phase.

One issue we need to be careful about is the spatiality of the sound, which can quickly get lost if we only record and replay sounds without any considerations about the physiology of the human ear and the relative position of the sound display device. As soon as the natural shape of the outer ear is circumvented, for example through the use of earbuds, sound spatiality can quickly become more shallow. Stereo mixing can still place a sound on a 2D plane around the user, but for higher precision or up-down localization we need to know some basic facts about the user's auditory system. Localization on the 2D plane can be much improved if we take into account head shape (or even upper-body shape) and specifically inter-ear distance, as the resolution of stereo signals is done depending on timing and phase differences in how a signal reaches each individual ear. Then for up-down localization these signals are encoded in complex ways in how the sound is modulated by the head and pinna. These basic physiological facts can be encoded in so-called head-related-transfer-functions (HRTFs) which describe how a sound that originates at a specific location will be modulated before it reaches the user's ear drums (Xie, 2013). There are some standardized HRTFs that approximate most human heads well, however full precision can only be reached through personalized transfer functions, usually based on a direct scan of the upper body.

2.2.1.3 Haptic Displays

Haptic displays are in many ways still an experimental category, however there is already a massive number of very distinct systems and concepts. It is far harder to categorize and group haptic displays into just two or three groups, as we did with visual and audio displays. This is because what we in spatial computing tend to simply refer to as "haptics" is actually a collection of sensations deriving from multiple sensory organs and physiological facts. There are tactile displays, usually serving their signal by direct skin-contact and mechanical vibration (Ishizuka and Miki, 2015) or through projection of air currents (Hoshi et al., 2010). There is force feedback, which selectively offers resistance to a user's movements to simulate digital objects and processes having a physical presence (Frisoli et al., 2005). And then there are related phenomena like pain reception and heat transfer, which are, with some exceptions (Saga, 2015), not often a part of display technologies.

Most of these system are wearables, like haptic gloves that can include both tactile feedback and force feedback to the fingers, exoskeletons that create force feedback for the whole body, or haptic vests that contain a multitude of motors to induce tacticle sensations over a large area of skin (Tanaka et al., 2002). A different class of haptic devices are tangible interfaces, which describe physically present data-objects that can be moved and manipulated by the user, usually to trigger reactions in a different object or in themselves, to display some sort of data (Fishkin, 2004). As such, they are the natural intersection of display and interaction device and stand in opposition to the long-lasting paradigm of generalized interface devices like mouse, keyboard, touchscreen and others. Many virtual environments that offer an immersive visualization try to emulate these kinds of data-objects by creating the illusion of haptic feedback through physicalized behaviours and audio-visual and sometimes tacticle feedback ("pseudo-haptics"), making their manipulation feel as natural as possible.

Just like see-through displays depend on advances in optics, the availability of better haptic displays depend on advances in non-computer science fields like material sciences and robotics.

2.2.1.4 Taste and Smell Displays

Much more experimental are the kinds of displays that induce taste and smell sensations, or even try to make use of some of the secondary senses in the human body. Taste and smell specifically are deeply entangled with each other, and very dependent on emotional and physiological states. More will be said about the peculiarities of these kinds of senses in Section 3.1, however some recent examples of displays that work towards them are: electrode-induced direct stimulation (Cheok and Karunanayaka, 2018), odor dispensers in static or wearable format (Nakamoto et al., 2009), odor beam forming (Hasegawa et al., 2018), edible data objects Mueller et al. (2021) or even tracked, edible objects that do not have much taste themselves, but like pseudo-haptics are given a form through other sensory impressions in order to trick the user's mind into tasting flavors that are not actually present (Narumi et al., 2011).

2.2.2 Tracking Technologies

Before any display can work or any interaction can happen, the software that orchestrates the virtual environment needs to know where the user and any other involved objects are located in space. Tracking systems are highly dependent on specialized hardware and calibration procedures, making their implementations particularly varied between separate manufacturers. There are however some broad categorizations we can make, based on the spatial relation between tracking system and tracked objects, the kinds of physical phenomenon used for tracking, what part of an object is tracked, the tracking scale, and the specific use case of the tracking system. Tracking is not as important as display or interaction for the purposes of this thesis, but there should still be a basic understanding of its realities.

First, the location of the tracking system can be separated into outside-in and inside-out tracking systems. Outside-in systems are installed in the environment, at a place where it can sense the tracked objects without interfering with them. Inside-out systems are installed on some or all of the tracked objects and track them in relation to the environment.

Secondly, there are different kinds of physical phenomena we can utilize for tracking (Bhatnagar, 1993). These include: Magnetic, electromagnetic, acoustic, mechanical, and inertial. Thirdly, there are tracking systems that track the environment as it is (markerless) and systems that need special tracking objects in the environment and on objects (marker-based). Tracking systems also differ in scale. They can reach from a system that tracks a single object at short distances (like a VR-HMD inside-out tracking a handheld controller) to Global Navigation Satellite Systems (GNSS). Lastly, systems differ by what objects they are supposed to track. Usually the focus are moving objects that are either static or have moving parts themselves, but perhaps the most interesting class of tracking systems in spatial computing are those that track parts and aspects of the human body, like gait or hand configuration.

2.2.3 Interaction Technologies

The way in which we interact with the world can be seen as just as important as the act of perceiving it. Interactions can reach from movements as simple as turning one's head while wearing a HMD, to complex series of learned movements that trigger a specific action within the inhabited world.

Usually interaction will require tracked devices, however with the growing capabilities of tracking systems, device-less body-driven interactions are becoming more attainable. In this way, our body is the most immediately available "interaction device" and allows us to skip any training phase we would have with a specialized device. This low barrier-of-entry is often taken and very uncritically translated into claims of body-tracked interactions being inherently more "intuitive" or "natural". This is questionable, as by this standard an arbitrarily complex language of learned hand signs would classify as more "natural" than an interface in which a stick-like interaction device is used to drive a virtual nail into a virtual piece of wood. Instead, these notions of intuitiveness are highly contextual—and intuitive often just derives from familiarity.

For a categorization, we first have to consider two different parts of each interaction: which action is being performed and which device or body part is used to enable us to perform this interaction. This will be discussed in detail in Section 3.3, by adapting Roth (2013a)'s framework of cartographic interaction to encompass both of these aspects.

Here, we will only introduce the different kinds of devices that are available. Interaction devices can broadly be categorized into:

Body Usually one of the limbs of the user, or another trackable biological function like speech, clapping, muscle movements or eye movements. Could also include prosthetics and other artificial body parts.

Held Usually meaning hand-held, i.e. a device that is light enough to be wielded and that can be quickly picked up or dropped.

Worn A device that can be temporarily affixed to the body, does not have to be grasped and is light enough to be carried. These devices usually can not be quickly dropped or picked up and need some time to be put on.

Static Immobile surfaces and objects that are located in the tracked environment.

Moving Objects and surfaces in the environment that can move on their own.

Moveable Objects in the environment that can be pushed, pulled, carried or otherwise moved by users.

Sometimes we of course can combine these categories, for example if a moving robot is light enough to be picked up and can be interacted with by the user. The standard assumption is also that these devices need to be tracked, though that is not always the case—controllers or keyboards can be moved and used without needing a physical presence in the virtual environment. Within these categories, there is also the question of whether a device is general-purpose or not. Handheld controllers for example commonly allow interactions that can be mapped to all kinds of applications (button presses, joystick movements), while moveable objects in the environment are often specific to one application.

The general-purpose handheld controller is probably the most common interaction device in spatial computing. Several interaction paradigms of such controllers are well studied, for example pointing interactions (Christou et al., 2016), virtual keyboards (Dudley et al., 2019) and grasping interactions (Chessa et al., 2019).

New interaction devices and paradigms in spatial computing are a massive research topic. The breadth of technical developments that are currently made are beyond the scope of this thesis. When considering the taxonomy of interaction in Section 3.3, we will instead stick to the device categories as shown above.

2.2.4 Immersion, Presence and Embodiment

After considering all of these technologies, we of course need to consider the motivation that stands behind using each of them. Every spatial computing system has an application area, some of which we will show in the next section—but between the application area and the choice for a spatial computing-based solution stands the question of what the motivation is for not simply using conventional, well-established, 2D means of computation, and instead going for difficult-to-use, largely unproven hard- and software. Ultimately, many of the commonly stated driving factors of spatial computing, like "immersion", are not strictly exclusive to spatial computing technologies. A well-made, interactive 2D map about an emotionally charged topic could be considered just as "immersive" and can easily run on the devices everyone has at home or in their pocket.

So why go spatial at all? When arguing for the usefulness of such technologies, two terms are very commonly invoked in combination with immersion: presence and embodiment. While *immersion* in Milgram's definition (Milgram and Kishino, 1994) refers to the technologically-enabled level of sensory replacement and can be objectively assessed through measures like field-of-view, display quality, affected senses, and more, the way the term is often used in common parlance is scientifically more accurately described by the concept of *presence* as defined by Slater and Wilbur (1997), which describes the feeling of "being there", i.e. the purely subjective responses to the virtual environment that technology immerses us in. More technically, Slater later describes it as "the propensity of people to respond to virtually generated sensory data as if they were real" (Slater, 2007). There is much debate about this term and its use, both in scientific papers as well as in the larger Spatial Computing community, which involves psychologists, developers, designers, artists and many more. On example would be popular podcast host Kent Bye and his *Elemental* theory of presence, in which he uses the evocative concept of four elements to more intuitively describe the sorts of immersion humans can feel—fire for "active presence", air for "social presence", water for "emotional presence" and earth for "embodied presence" Bye (2017). While these more creatively driven definitions are being formulated, Mel Slater has long been trying to establish more rigorous and clear definitions, for example in his 2007 lecture notes (Slater, 2007), in which he makes it explicit that not all subjective impressions are presence and that the term must be separated from, for example, sensory concepts like metamerism, in which different wavelength distributions in visible light appear as the same colors to us, or emotional concepts like *involvement*, which separates our feeling of being present in a space from our emotional engagement with it. The latter is quite similar to the notion of water, i.e. emotional presence in Bye's model—which illustrates the issue that these terms are so often confounded, that we have no choice but to pick a definition for our current application and try to be internally consistent throughout.

Ultimately, immersion and presence are very important for spatial computing as a whole, but only of secondary importance for this thesis. Visualizations only sometimes require the feeling of presence, and specific measures of immersion enabled by a specific set of technologies rarely become important. Much more central is one of the concepts often related or confounded with these two terms: *Embodiment*. Like all the other common words in this space, we have to grapple with multiple definitions here. First there are definitions of embodiment that are not even close to the spatial computing discipline. Most basic and applicable here perhaps is the simple user-experience focused definition from Hartson and Pyla (2018), which defines embodied interactions as "Interactions with technology that involve a user's body in a natural and significant way, such as by using gestures". In Slater's lecture notes "embodied" appears as meaning that a user is shown a simulated virtual body in the virtual environment that tries to match the movements of their real body, which has overlap with a long standing discussion in the VR space around whether only body-parts that are actively tracked should be shown, for example as ghostly "floating hands", or whether an attempt should be made to show a full body even with imperfect tracking. A similar notion of embodiment also appears in Ahn et al. (2016), where the user is given an animal body in the virtual environment to increase engagement with the displayed ecological issues. Chen et al. (2017) instead take the term quite literally and create a collaborative environment in which a group of participants each use their body to play the role of one data point in a data set. Fishkin (2004) is closer to Hartson and Pyla (2018)'s definition, however it flips what body is in focus: their embodiment describes the feeling of of how much a computation (in our case data analysis and representation) is embodied within the current state of a haptic display device. Liu et al. (2020) apply the concept of embodiment explicitly to geovisualization but focus on the importance of multimodal (i.e. spanning multiple sensory modalities) display and interaction in such systems.

For one definition far outside of spatial computing, we have the concept of "Embodied Mapping", in which maps are related to the body in a performative way, such as historical art in which the map of a country is imprinted upon the likeness of its ruler (Perkins, 2009). Many such definitions exist in the humanities, however they are beyond the scope of what we aim for here.

Then there is the main reason why the term of embodiment has become so popular in recent times: cognitive science and philosophy. In the cognitive sciences, a recent, non-classical approach to human cognition is the theory of "Embodied Cognition", described at length by philosophy professor Lawrence Shapiro in Shapiro (2010). Embodied Cognition opposes the view that cognition is based on the human brain operating as a computer-like symbol-processor (symbols in this case being multisensory impressions of the world) which receives its inputs from the sensory systems of the body, and instead describes cognition as a highly contingent, body-wide process in which the brain is simply one of many participants in a sensory-cognitive feedback loop. This is to some degree seen as obvious by most scholars in the field, however the specific reasoning and conclusions of modern embodied cognition theory are hotly contested, both in the cognitive sciences, as well as in philosophy, often debated as one part of the "4E's" of cognition—embodied, embedded, enactive, extended (Maiese, 2018; Shapiro, 2010; Newen et al., 2018). This debate goes far beyond the scope of this thesis, however the mental models we use do not always have to be entirely correct in order to be useful in the circumstances that we want to apply them in. For now, we can simply take these concepts as the basis for a conjecture: Engaging with data through visualization is a cognitive process. Given that cognition is a body-wide process, it can be useful to involve the full body in data analysis.

Considering all these definitions, what will embodiment refer to when invoking the term during the rest of this thesis? Because most of our focus is on display and interaction technologies and the combination of them, it will refer to Hartson and Pyla (2018)'s definition of interaction technologies "that involve the user's body in a natural and significant way", as well as to the sensory-cognitive process of perceiving and understanding data over multiple sensory modalities.

Definition 1. *A system of display and interaction technologies is **Embodied** if its use requires significant and natural body motions and if it implements a body-wide sensory-cognitive feedback loop.*

2.2.5 Spatial Computing Applications

2.2.5.1 Virtual Environments

Spatial computing applications, even if they are not geospatial, will always create or refer to some sort of virtual environment. When we directly use an existing, i.e. real, environment as-is, we tend to call these applications *location-based services*. If a location-based service and spatial computing are combined, all added virtual elements are embedded into the real location and are in direct reference to it. Sometimes interactions with these virtual elements can even have a direct effect on the real environment. Depending on the location we refer to, this often overlaps or integrates with related concepts such as the internet of things (for example in home automation or sensor networks) (Kamilaris and Ostermann, 2018) or "digital twins" (for example in manufacturing and urban planning) (Hasan et al., 2022).

We can also use existing environments more indirectly by fully turning them into *scanned* virtual environments through different scanning technologies. On-the-fly scanning of environments often already lies at the heart of tracking algorithms like visual simultaneous localization and tracking (SLAM) (Taketomi et al., 2017), so this method is sometimes directly related to location-based services. In scanned virtual environments we however use the scan itself for direct display and interaction, sometimes even changing the environment from its original state. Common application of such environments are laser-scanned buildings and interior spaces Dinis et al. (2020).

If we do not want to refer to real-world directly at all, we can create virtual environments entirely from scratch, usually in the form manually or procedurally generated textured polygonal models. In this thesis this will be referred to as *simulated* environments. These environments tend to offer the highest degree of interactivity. They can further be divided into environments that try to appear as realistic as possible, usually in the hopes to induce a high degree of presence, or environments that are focused on one specific use. In these latter environments the interaction and its result are the point, and they can operate under entirely different mechanics than the real world. Such environments are not always entirely free of real world influence however, as an environment that was procedurally created based on some spatially limited information (e.g. processed, lower-resolution spatial data instead of a direct optical scan) about a real place could still fall under this category.

2.2.5.2 Application Domains

Another, more obvious categorization is of course a categorization by the domain and goal of the application. This is also much broader than the categorization by used environments, so only a small cross-section will be discussed here.

Most immediately adjacent to the geospatial focus of this thesis is the field of surveying as well as any kind of field work that involves spatial data. Geospatial data is in many ways a hidden layer with information about the real world, however it is still difficult to access this data while away from a computer. Spatial computing, if used correctly, enables us to access that hidden layer and reproject it onto our surroundings, aiding us in further data acquisition or other forms of spatial reasoning. This is no easy feat however, as the sort of tracking usually employed for augmented reality is not made for the sorts of sub-centimeter accuracies that are common in surveying, to not even speak of the geospatial registration of locally tracked surroundings with the wider geospatial reference systems. Seminal work in this area has been done by Kreuziger et al. (2015), however the problem is still far from being solved. Despite the accuracy issues, there is research into how to actually display this geospatial information, all under the assumption that one day we will have something close to on-demand global sub-meter accuracy. Some of the most important work in this area originates from Langlotz et al. (2014) as well as Zollmann et al. (2014) and Grubert et al. (2016), whose work revolves around superimposing common geospatial data like building labels, street names or even sub-surface data into the user's field of vision. It is perhaps surprising that basic, mostly answered questions of cartography like label placement suddenly start to reappear when we are adopting an egocentric viewpoint within the real environment (Postert et al., 2021).

As we move further away from the domain of this thesis, we find what is perhaps the most prominent application domain of all: *teaching* and *training*. Simulated virtual environments where haptic controllers are paired with visual and audio hints as to how to operate a machine, tool or vehicle, if used right, have the potential to be highly effective (Howard et al., 2021; Cheiran et al., 2020) in developing familiarity and muscle memory where training on the actual object is too expensive or dangerous, like for example in oil rigs (Wan et al., 2020). Studies have also found large effects in utilizing virtual environments in the teaching of spatial concepts in science, technology, engineering and mathematics (STEM) education (Merchant et al., 2013; Lee-Cultura, 2019; Safadel and White, 2020), as well as in group-based learning (Souza et al., 2020). With regards to the geospatial focus of this thesis, there are also research results showing how to use immersive geovisualization in education (Philips et al., 2015), for example through tangible interfaces like sand desks for remote sensing education (O'Banion et al., 2022) and fully integrated learning systems that incorporate drone-based image acquisition (Bernardes et al., 2018). Another relevant note for this thesis is the potential for multi-sensory applications of learning. Research here searches for the potentials of engaging all the body's senses to increase learning outcomes (Shams and Seitz, 2008; Roberts and Roberts, 2014), however quantifiable results still remain somewhat sparse.

A related domain is *medicine*, where spatial computing interfaces are often used for training as well. One of the more prominent areas of research here is surgery education (Rogers et al., 2021; Ayoub and Pulijala, 2019), which is also one of the application domains where haptic interfaces are most immediately useful, see for example Medellin-Castillo et al. (2021) and Imran et al. (2021). Beyond just training and education, there is research into AR assistance interfaces (Cutolo et al., 2019) and even telemedicine (Huang et al., 2019). There also has been progress in therapeutic treatments for neurological damage (Matamala-Gomez et al., 2021) and anxiety disorders (Boeldt et al., 2019).

Another obvious application that has long been a dream of many researchers and corporations is VR *teleconferencing* (Greenhalgh and Benford, 1995). There is still current research into this topic (Campbell et al., 2019), and the technology has already been applied in many circumstances, especially during the COVID-19 pandemic. Recent hype around the controversial concept of the *metaverse* (Mystakidis, 2022) has shifted expectations in this space from conferencing in small groups to massive multi-user environments.

One application type that encompasses many domains is the use for spatial computing as a *presentation* medium. This could be advertising (presenting a product) (Feng and Mueller, 2019; Kim, 2021), virtual tours (presenting a place) (Beck et al., 2019; Westphal and Berger, 2020) or simply just presentation of ideas (where presenting overlaps with education) (Boetje and van Ginkel, 2021; van Ginkel et al., 2019). Virtual tours especially might face a future trend where they do not show existing places, but allow us to record and relive spaces that are long gone in a very visceral fashion. A glimpse into what this could do is offered by the field of virtual heritage, where approximations of ancient sites or still-standing cultural heritage sites are scanned and then enriched as virtual environments (Gaitatzes et al., 2001).

A subset of most of these domains is the deployment of spatial computing for *collaboration*. Any virtual environment could potentially be turned into a multi-user environment, where you could learn (Jackson and Fagan, 2000; Jochecová et al., 2022) and train together with others (Khanal et al., 2014), plan product designs (Shen et al., 2010) or conduct team-based task (Kockro et al., 2007) with other domain experts.

Perhaps the most exciting, but also the most nebulous application field is the invention of entirely novel virtual tools. The non-physical, consequence-free nature of virtual environments allows us to employ interaction objects and physics that would not be possible, or far too dangerous, in the real world. Recent examples are a new approach to interacting with spectrograms for audio design in Engeln et al. (2018) or Crawford (2019)'s "hyperphysical" interfaces, which inspired the work in Chapter 5.4 of this thesis. These interfaces take the notion of embodiment and use it to extend the user's body with playful new interaction affordances, from entirely new limbs to non-standard physical interactions with existing ones. In the best possible case, such tools would feel like “superpowers” in what they enable us to do—a concept we will return to in Section 4.3 of this thesis. Tools that reach such a level of quality will of course be challenging to develop—but they could be transformative enough that the effort is worth it.

Finally, one application domain that deserves specific note in this thesis is the use of spatial computing in architecture and civil engineering. Next to the already noted domains of gaming, medicine, and training, this is perhaps one of the most popular application areas in recent years, as the whole field is shifting to high-quality 3D models that can be used for advertisement and planning. It also has special relevance to this thesis, because it is highly entangled with the geospatial sciences. Both fields share many of the same technologies and data analysis issues. Possible applications include building design through data visualization and immersive 3D model-based simulations (Zaker and Coloma, 2018; Safikhani et al., 2022), architecture education (Jensen, 2017) and public participation (Chowdhury and Schnabel, 2019).

2.3 Immersive Geovisualization

Now that we have established a solid foundation in both geovisualization as well as spatial computing interfaces, the next step is to combine the two. Just like its two parent fields, the field of immersive geovisualization already has a varied history, with even its name going back as far as the late 1990s (MacEachren et al., 1999). But as with all spatial computing-related research, the recent resurgence of off-the-shelf consumer hardware has infused new life into this field, resulting in a wave of new research both for immersive geovisualization specifically, as well as for the larger field of immersive analytics. Because of how varied this research is, this section will be as much about what this thesis will *not* focus on or attempt, than about what is actually in focus.

There are many different approaches to this kind of research. Some studies try to "simply" translate visualization techniques like multiple coordinated views directly into the immersive domain (Mahmood et al., 2018) or directly compare the performance of certain techniques in immersive analytics to their non-immersive versions (Sardana et al., 2021). Other papers directly criticise this approach, instead recommending to engage immersive analytics in its own *embodied* way, while already attempting to create conceptual models for developing well functioning immersive visualizations (Gračanin, 2018).

Even the recent work is still very foundational. Perhaps the most important publication in the space is the book "Immersive Analytics" by Marriott et al. (2018), which collected most of the existing research and research directions in immersive analytics into one place and identified many different research questions for the coming years. Notably, while many papers and books about visualization research include a special section for cartography and geovisualization, Marriott et al. make no such difference. The sections most specifically related to spatial data are about "Situated Analytics" (more on this later) and "Exploring Immersive Analytics for the Built Environment", but even then the geospatial is just one aspect of many.

Apart from technical considerations, perhaps the biggest question of the field, as posed in Ens et al. (2022), is how "to define which studies need to be conducted to assess the effect of embodied interaction on cognition in data analytics." Spatial computing research offers many, at least preliminary, answers about certain effects that immersion and embodiment have on our cognition and our bodies, however there is only limited research about the specific influence they have on our ability to reason about data, especially in comparison to traditional visualization systems (Ens et al., 2022).

In the sub-field of immersive geovisualization specifically, there is a large amount of interest in how we can or should conceive of immersive environments when visualising geospatial data. Hruby et al. (2019a) for example take a technical look at how to design immersive geovisualization that maximize presence, considering questions like how to properly move from 2D GIS data to 3D models fit for 1:1 scale display and where to introduce dynamic vs. static elements. This is validated with the example of a tool for VR coral reef exploration. Here, immersive geovisualization is conceptualized as something that happens by creating a realistic and immersive virtual environment.

Many publications in this field also deal specifically with the differences between conventional cartography and immersive geospatial environments. The main factors that current papers look at here are *scale*, i.e. specifically the focus on the 1:1 scale and its implications as opposed to normal cartographic scales (Hruby, 2019a; Zhao et al., 2020; Hruby et al., 2021), *presence* and *immersion*, what they do and how effective they are for geographic environments and data sets (Hruby et al., 2020; Rzeszewski and Naji, 2022; Newell et al., 2017), and collaborative environments for visualization and data analytics (Dolezal et al., 2020; Lee et al., 2020; Chen et al., 2021). There are also debates about whether immersive geovisualization as a project is worth it at all, as immersive geovisualization takes all the problems that are associated with 3D geovisualizations and brings them much closer to the user, with much more uncomfortable and bulky hardware, adding many physiological and perceptual troubles into the mix. Most usability studies at the moment seem to suggest distinct positive and negative effects, like cognitive processing of global scenes being improved while readability of local details is reduced (Çöltekin et al., 2016). However, these effects seem to be highly dependent on sociodemographic differences in the users (Chassin et al., 2022). We can likely expect higher effect sizes and clearer results as the technology matures further and more people attain some level of familiarity with it.

Moving on from the differences to conventional systems, what follows will be an overview of domains where immersive geovisualization has been used. Lorenz et al. (2008), Veas et al. (2012) and more recently Spur et al. (2022) create multi-perspective views, where despite the 1:1 egocentric view the terrain around the user is curved upward such that all the data can be seen from every point in the data set. This thesis will return to this in detail in Section 5.4. Zhang et al. (2021) pairs immersive analytics with urban site planning by including a heavy focus on object placement interactions. Spur et al. (2020) implement immersive viewing of and physical interactions with multiple map layers at once and investigate the effects this has on assessing public lighting maps. Wagner Filho et al. (2019) adapt the common spatiotemporal visualization technique of the "space-time cube" (for example seen in Andrienko et al. (2014)) into an immersive environment and assess its effectiveness. Rydvanskiy and Hedley (2021) investigate the current state of mixed reality flood visualization—a commonly researched use-case for such systems, as they could allow us to viscerally see what floodings would do to our immediate surroundings. Klaas and Roopaei (2021) surveys the application of immersive analytics for herd behaviour and herd monitoring in smart agriculture. Cartwright et al. (2022) tries to use the advantages of immersive analytics for water reservoir engineering, as there are some complex visualization situations involving 3D subsurface data sets that could benefit from new approaches.

Applications that do not involve geospatial data can still contain spatial interactions with data that has an inherent geometry, such as in the immersive lenses project (Kluge et al., 2019), where virtual embodied tools are used to investigate the inner contents of volumetric data sets such as 3D sonar scans. The data can also be fully abstract, as is the case in a lot of the foundational research in the field of immersive analytics, for example the investigation into 3D scatterplots and with which interaction paradigm to navigate them in Yang et al. (2020a). In the field of medicine, there is the VROOM research project by Lau et al. (2022), which creates an abstract immersive environment for oncology analytics. One of the most impactful research projects in the space is the "ImAxes" research in Cordeil et al. (2017), where a completely novel interface for exploring abstract data within consumer VR hardware is proposed, implemented and released. Users can move data around with tracked motion controllers and through the movements and placement of different interaction elements trigger various visualizations, thus creating a fully embodied visualization platform. ImAxes has already found use in geovisualization, as it was adapted by Cunningham et al. (2021) for embodied energy sector analytics and by Newbury et al. (2021) for immersive, spatialized map layers.

Before we move on to multisensory and embodied applications specifically, we need to look in more detail at the aspect of *environments* in immersive geovisualization. This will allows us to establish clear boundaries on what kinds of data the systems considered in this thesis are supposed to display—which will constrain what multisensory representations are important to us.

2.3.1 Virtual Geographic Environments

The displayed environmental context is one of the most important aspects of any immersive geovisualization. This was already noted before when considering scale, object placement and required degree of realism, but deserves more specific focus.

The most direct way to define these environments would be to adapt the term of the virtual environment from Section 2.2.5.1, and then simply specify that the environment is geographic in nature, i.e. is referencing some real place and tries to represent it to some degree of immersion. A common definition and conception for such virtual geographic environments (VGE) can for example be found in (Lin et al., 2013). Broadly, they are environments that display and allow spatial (and sometimes analytical) exploration of geographic facts. These environments are usually built in a 3D game engine like Unity or Unreal Engine, which are repurposed as visualization tools, where the mechanisms of game object placement and shaders are used to semi-procedurally build an interactive world, often based on combinations of multiple geospatial data sets like digital elevation and city models (Keil et al., 2021).

One of the main disconnects between VGEs and current GIS use cases is that the conventional layer-based or feature-based data representation methods are often not of high enough fidelity for direct immersive display (Hruby et al., 2019a). One way to approach this is to apply a number of manual and automatic steps to move from GIS data *features* to game engine *objects*, such as shown in Hruby et al. (2019a). Such objects could be animated or even have limited forms of agency in order to increase the realism of the virtual environment. In Hruby et al. (2019b) these environments are not conceptualized through the very generalized (and often quite nebulous) concept of a VGE, but as “geovisualization immersive environments” (GeoIVE), which have five defining criteria:

1. Interactive exploration is a must to induce spatial presence, as is a stereoscopic vision either induced through a HMD, stereo glasses, or an autostereoscopic display.

2. The immersive virtual environment must be modeled and moved through at a 1:1 scale in relation to the real world data.

3. The 1:1 scale implies a sufficient level of detail that allows us to distinguish and locate specific real-world objects that are present in the virtual environment.

4. The environment must be realistic not just in scale, but also in appearance.

5. The environment is a model, and as such only the interesting or relevant parts of it must be modeled at high fidelity.

The main focus in such systems is to induce (spatial) presence, to get users mentally situated into a virtual environment and make them engage with it fully. This allows users to viscerally come into contact with natural phenomena that they normally would only experience in an abstract way, like ocean acidification. This explains points 2 and 4 respectively, which might otherwise be controversial on first glance. After all, 1:1 scale must not always be retained if data visualization is the goal, and in some cases might even be a hindrance. The popular Google Earth VR application shows that it can be a powerful tool to be able to change environmental scale on the fly in virtual environments (Käser et al., 2017). Similarly, realism of representation is not as important for an application focused on communicating the statistical attribute values encoded in geospatial data. Invoking spatial presence is a different process than conventional cartographic representation (Hruby et al., 2021)—in fact, the complexities of representing a realistic environment could sometimes be an active hindrance in transporting geostatistical results with clarity.

As such we can distinguish the type of immersive geovisualization we are aiming at in this thesis as something more akin to immersive analytics as defined in Marriott et al. (2018), but applied to geospatial data. This visualization *can* occur in GeoIVE, but does not have to. The role of the environment is not necessarily to generate spatial presence, but to establish a geospatial reference and to enable data cognition in immersive and embodied ways. Of the five criteria however, at least 1 and 5 arguably still hold true. In many cases, criterion 3 will also remain important. The data might not have to be visualized at a 1:1 scale anymore, but it still needs to hold up in a context in which a user has full control of their viewpoint and expects to be able to interact with data in deep, embodied ways.

Another way to look at immersive environments for geovisualization is an older categorization from Hugues et al. (2011). Here, in a pun on the now-famous saying by philosopher Alfred Korzybski that the "map is not the territory", the authors distinguish between *augmented maps* and *augmented territory*. Augmented maps are maps represented as 3D digital models that are then put into the user's environment (the paper considers only the AR use case, however the principles remain the same for different forms of immersion) as a virtual object or full virtual environment to explore. Augmented territories are the addition of virtual objects in direct reference to the real environment. This can be a very useful classification to use for the hardware- and paradigm-agnostic approach taken in this thesis, as otherwise certain applications are often conflated with the display technologies commonly utilized for them.

Another way to consider environments can be found in Lü et al. (2018), where the authors postulate the concept of the *scene*, or *geographic scenario*. Here, it is not the realistic representation that is in focus but the dynamic and accurate simulation of processes underlying a visualized geographic phenomenon, so that they can be interacted with in deep and integrated ways. Because of their more physical, simulated character, such environments tend to invite further deliberation on data aspects that are often missed in conventional maps and GIS visualizations, like evolutionary processes, accuracy measures and the semantics of interactions between different data sets (Chen and Lin, 2018). This makes the concept of the geographic *scene* the most appropriate way to conceptualize the environments in this thesis. The levels of immersion and presence can vary—the interactivity and dynamics of the scene are the point. Use cases include urban modeling and participatory planning (Lin et al., 2022; Chassin et al., 2022), dam failure simulation (Yu et al., 2021), debris flow disasters (Li et al., 2022), water pollution control (Rink et al., 2018), crowd evacuation (Song et al., 2013), and more. The demands that such scenarios place on data representation are one of the factors that will motivate the taxonomy of spatiotemporal data types in Section 3.2.

To summarize, realism of the rendered environment is often considered the main advantage of immersive geovisualization system. However, because we are focusing on analytical sense-making akin to visual analytics, realism becomes a secondary consideration. What is more important are the kinds of interactivity that are enabled by a geograpic scene, as well as the way it is referenced to the local tracking space. To conceptualize the latter aspect, we can look back to Section 2.2.5.1: here, environments were treated as either simulated, scanned, or location-based. For the geospatial case that means the following:

Simulated A simulated scene is one in which geospatial data is manually or procedurally enhanced into a representation that mirrors certain aspects of the real environment.

Scanned A scanned scene is one in which a real world environment is scanned directly and then processed to be an operational interactive scene.

Location-based A location based scene is one in which the user is physically located near the data, and only the relevant scene objects need to be digitally represented.

This categorization is one of the concepts that will be used in Chapter 4 to give environmental context to to the model we are developing. What is important for the remainder of this section, is that while simulated and scanned scenes behave similarly in many ways, location-based scenes have many additional complexities. The research branch that is specifically concerned with them is called "situated analytics".

2.3.2 Situated Analytics

Situated analytics as defined by Marriott et al. (2018) is a method that "employs data representations organized in relation to germane objects, places, and persons for the purpose of understanding, sensemaking, and decision-making." The relationship to spatial computing and geospatial data should be clear—to even enable this, the whole breadth of spatial computing technologies is necessary, as well as a solid, geolocated data set fit to be utilized in a space. The payoff for such a system however could be immense, as they allow us to visualize and analyse in those moments when we are spatially and temporally most closely located to a phenomenon or site, and thus most able to intervene or act upon it.

Examples for this sort of system already appeared earlier in this thesis, for example the geodata browser work of Zollmann et al. (2014) and Grubert et al. (2016). Another example is work by Guarese et al. (2020), in which an AR system is employed which allows users to analyze and make decisions about their actions in indoor spaces such as auditoriums or classroom, for example with respect to accessibility (both wheelchair and hearing-related) of certain seating situations, wifi signal strength, air flow, and exit access. A more geospatially-focused application can be found in White and Feiner (2009)'s SiteLens project, which looks at different symbologies for displaying air quality data at street level.

Most of the current examples however are not very complete and are most concerned with investigating specific aspects of the tracking, interaction or display that such systems could soon employ. While doing so, they always have to grapple with inadequate display technologies, issues with localization accuracy, as well as data resolution that is not sufficient for accurate real-world placement.

Despite the distinct lack of well integrated examples, Marriott et al. (2018) attempt to conceptualize many of the aspects of situated analytics. They introduce the axes of *situatedness* and *analytical level*, where situated analytics systems are both of a high situatedness and high analytical level, while simple ambient information displays or untracked data glasses without content registration are of a low situatedness and low analytical level. The concept of situatedness has conceptual overlap with our concepts of simulated, scanned and location-based scenes, and can serve as an appropriate title for this category.

Marriott et al. (2018) also relate situated analytics to many other fields like ubiquitous computing, embedded data representations, and of course augmented reality. They also correctly note that many data sets are not just geospatial, but spatio-temporal and as such situated analytics can also be situated in a specific time frame.

Included in this analysis of situated analytics is a conceptual model for how to apply it. This includes a new situated visualization pipeline which includes concepts like physical referent (the real object that some data references) and the physical representations (that make the visualization visible to the user). However, this model is too far removed from geospatial data to be of much use here: Its focus lies on object-based situated analytics, in which objects in the environment are extended with analytics visualizations. This might be a useful lens in certain geovisualizations with tangible objects, but will not be in focus here, as Section 4.3 will show a more fitting combination of models that generalize better over situated and non-situated system for the geospatial case.

2.3.3 Multisensory and Embodied Analytics

With most of the trends and terms we will need now defined, we arrive at the last missing piece: If we want true embodiment, i.e. to truly involve the whole body in our data analysis, then we need to speak to more than just the visual sense. A true *Embodied Analytics* system has to be immersive for more than one sense and interactive for more than just our hands.

An interesting aspect of the concept of multi-modal representation is that the incongruent term of "multisensory visualization" has caught on in some publications. This is easy to explain, as we have learned the term visualization carries with it much more context than the simple fact of visual representation. Still, to be specific at least in this thesis, from here on out any reference to data representations that are not specifically visual will be made with the term *sensification*, adopted from a paper formulating research directions in multisensory data representation by Tak and Toet (2013). This term neatly serves as a parent term for so many of the kinds of representation in the literature, from visualization, over sonification through to physicalization. A similar term that appears in the literature is *sensualization* (Ogi and Hirose, 1997).

Definition 2. *A **sensification** is an interactive physical or digital representation of a piece of data or concept to multiple human senses at once.*

The question now is: has there already been research into embodied immersive sensifications for the purposes of data exploration and analysis, specifically those that incorporate multiple senses at once? As before, the place to start with this question is Marriott et al. (2018), which devote an entire chapter to "Multisensory Immersive Analytics". The chapter contains a surprising number of examples, conceptualizations and physiological facts about the human sensory system as related to data sensification. Most crucial is the proposed conceptual model for multisensorial immersive analytics, which starts with an extended, multi-sensory visualization pipeline (see Figure 2.3).

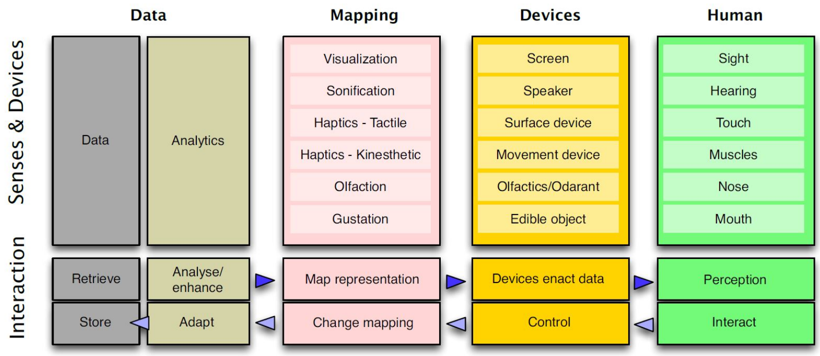


FIGURE 2.3: The multisensory analytics pipeline, showing the feedback loop between data, display devices and human, from Marriott et al. (2018).

This new pipeline makes explicit what in the original visualization pipeline was implicit: in order to perceive the data, the display device needs to enact the data mapping onto a human sensory organ. It also makes explicit the feedback loop inherent in any interactive visualization: We perceive the representation, then we act on the representation through our control devices, in turn changing what we perceive.

With the pipeline, Marriott et al. also introduce the concept of sensory mappings: the mapping of data values to sensory variables, i.e. signals that can be sent through a display device and perceived by the human body. The pipeline and sensory mappings are the basis for a lot of the work in this thesis and thus will keep appearing throughout Chapters 3 and 4. What the chapter does not do, is to relate these new concepts to geospatial representations—and therein lies the problem: most existing research in this space is so foundational that it has only occasionally been applied to the complexities of geographic data. The chapter also focuses on display technologies, often omitting the role of interaction in these circumstances. This poses the question: What could multisensory, embodied *geo-sensifications* look like? (From here on simply referred to as *embodied geosensifications*.)

Many larger geovisualization review papers and research outlooks directly state that other display modalities should be investigated, but then go on to state that they are out-of-scope as of now, categorising them as future research possibilities (Lü et al., 2018; Çöltekin et al., 2020). This is surprising, as at least two sub-domains of this problem are fairly well explored: audiovisual cartography and haptic maps. Most of the research here of course concerns non-immersive systems, but the inclusion of immersive display and interaction hardware is not unheard of (Edler et al., 2019; Hruby, 2019b).

In audiovisual cartography, traditional maps or geovisualizations are enriched with different kinds of sounds that can display additional data, increase the sense of presence, or represent a real sound situation. The combination of these sounds is called the *soundscape*, in reference to the landscape that the viewer watches or inhabits. Edler et al. (2019) distinguish three different kinds of sound: abstract sounds, speech, and music. Each type of sound can be applied in different ways—speech could be created by virtual agents in the space (for example in simulated crowds) in order to increase presence, while music could be modulated to represent trends in one of the represented data sets. Our focus here of course is displaying data. When data is displayed through sonic means, we enter the domain of sonification, a long standing and in some aspects already mature discipline that grew next to and together with visualization (Dubus and Bresin, 2013). Sonification has been directly applied to geospatial data for use cases ranging from large scale coverage data sets (Schito and Fabrikant, 2018; Berger, 2020) to smaller scale display of noise data (Berger and Bill, 2019). There are also data-driven, but less goal-oriented examples, like the "urban musifications" in (Schetinger et al., 2021).

Haptic maps on the other hand sometimes appear on their own (Griffin, 2001), usually in the form of tangible objects (Petrasova et al., 2018) or as tracked sand tables (Harmon, 2016), but are more commonly associated with audio maps (Rice et al., 2005). This combination is especially common in maps for the visually impaired (Wang et al., 2009; Zeng et al., 2011). In these maps, sonic and tactile signals are put into carefully orchestrated interplay to allow the transport of spatial knowledge without any visual information being necessary. Haptic display techniques in general also appear in non-geospatial embodied immersive analytics tasks, for example when analysing volumetric scatterplot data as in Prouzeau et al. (2019).

Arguably, haptics also has to be involved when dealing with embodied visualizations, as at least the interaction component often utilizes haptic devices like controllers or tangible objects. An example is the *GeoGate* project shown in Ssin et al. (2019), which builds on ring-shaped focus and context techniques from Tominski et al. (2006) and Krüger et al. (2013) and embodies these techniques as a tangible ring device that is placed on a surface that serves as the interaction element for an augmented desk visualization.

Solid research for other senses like taste and smell is sparse to non-existent. For smell specifically there are some recent papers highlighting its effects on wayfinding (Schwarz and Hamburger, 2022; Hamburger and Knauff, 2019), which might in some way be exploited during multisensory data display, but is pretty far removed from these use cases overall. There is also research that is focused on the creation of immersive *places*, in which multi-sensory considerations are an important part Globa et al. (2022)—however these rarely use multi-sensory display systems beyond common visual, sonic and haptic interfaces.

The reasons for going beyond the visual can be manifold. Circumventing sensory issues in certain users is an obvious one. Less obvious are positive effects on certain aspects on cognition, like the inclusion of audio readings of location names improving spatial accuracy and memory tasks while reading maps (Lammert-Siepmann et al., 2017). There are many studies which cognitively investigate specific aspects of sensory effects in map reading and virtual environments (Siepmann et al., 2020; Rajguru et al., 2020), at least in the audiovisual case. Hruby (2019b) extends the concept of GeoIVE (see Section 2.3.1), to provide stimuli to multiple senses in order to increase the amount of spatial presence beyond what visual-only environments can create.

One of the largest works in creating a model for multisensory, immersive visualizations can be found in Moloney et al. (2018). The central claim here is that complex environments that present a dense amount of information over multiple senses can actually work better than the very focused "sterile" information visualizations we are often accustomed to, as long as they are tuned to the ecological perception of natural environments present in humans. Notably, these environments do not have to be generated through geospatial data. Moloney et al. start at the concept of *affordances*, which refers to "possibilities of action" that a certain technology, device or virtual tool enables for us, and ask the question: "what is the affordance of VR technology for the immersive analytics of big data?" This is in reference to a stream in HCI research, which focuses on an action-oriented model, in which users are not conceptualized as seeing the environment through their semantic knowledge and interact with it to receive information streams, but as "tool users" who latch onto the potentials for action that the environment *affords* to them.

Their work on a model starts at a basic level: they extend Milgrams's reality-virtuality continuum with the senses and the different levels of situatedness (see Figure 2.4) and the basic feedback loop of a knowledge generation model for visual analytics (see Figure 2.5). This knowledge generation model, originally from Sacha et al. (2014) is close to the sensory, embodied feedback loop that was alluded to before in this section, where the sensory feedback we get influences what we do next to change this same sensory feedback—but here the focus lies on the generation of hypothesis, insights and knowledge, and how they influence the actions we take within a virtual environment.

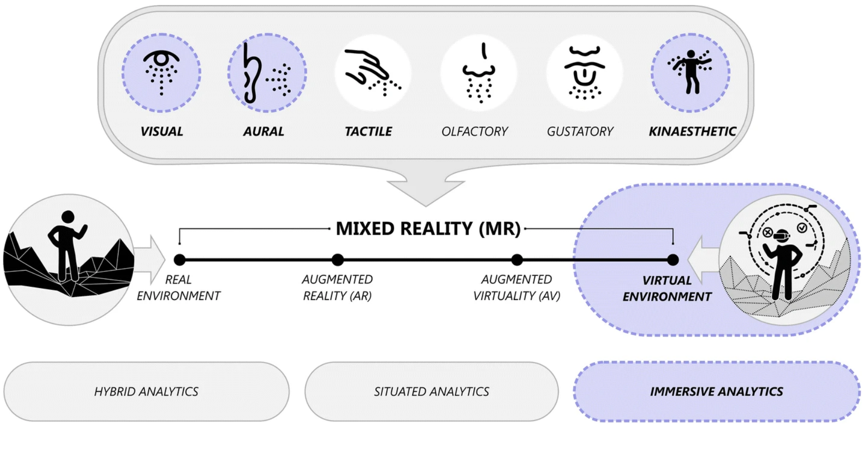


FIGURE 2.4: Milgram and Kishino (1994)'s reality continuum extended for multi-sensory analytics, from Moloney et al. (2018).

An interesting note is that in Figure 2.4, the haptic sense is split into tactile and kinesthetic. This is somewhat arbitrary and leaves out other aspect of haptics, like temperature. This view of sensory reality also translates to later considerations within their paper, where the sensory affordances they list (i.e. the variables that a virtual environment can use to transport information) are not ordered strictly by the senses, but under the categories of space, shape, colour, lighting, motion, physics, and aural. This distinction makes sense within the ecological scene-building context presented in the paper, but seen on its own commits several category errors, like distinguishing lighting and colour.

The paper also considers that for an ecologically valid virtual environment, i.e. an environment that our intuitions and natural predispositions can apply to, the sensory mappings we use for certain objects and data can not be entirely arbitrary. One example noted is that visually large objects should create different sounds than visually small objects. Our understanding of the specifics of this is still limited, but it is important to keep in mind when designing multisensory environments.

Moloney et al. go on to paint a very evocative picture of such ecological multisensory analytics. Several principles of how data should be represented and encoded are introduced. Ultimately, this conception is quite different from the one in this thesis, because it *spatializes* arbitrary data, while we are dealing with data that

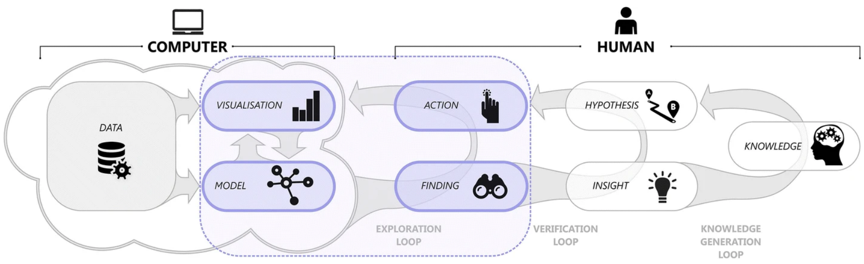


FIGURE 2.5: The knowledge generation model for visual analytics, originally from Sacha et al. (2014), redrawn in Moloney et al. (2018).

is already spatial. We do not encode data into environments—the environment is the data. We could theoretically create an ecological presentation of forest data set, and represent it as a city, in which the health of each tree is represented by how derelict each building is. The usefulness of such an endeavour however is questionable. And while some of the principles might be useful here—like salient parts of the data being highlighted to multiple senses in multiple ways—we will break others. One example is the principle that data representations should follow natural precedent. Certain aspects of our geosensifications can be specifically *without* natural precedent, in order to find new ways to conceive of problems outside of the constraints of nature (as will be shown in Section 5.4).

2.3.4 Embodied Geosensification

At the end of this chapter, there remains the question if there are any working systems that bring all these different factors together and that can already be found in the literature. First, we will start with a clear definition.

Definition 3. ***Embodied geosensifications*** *are immersive sensifications whose interactions are embodied to a significant degree and which facilitate* exploration of *and* analytical reasoning about *geospatial features within a geographic scene.*

Many publications implement systems that fulfill some of the requirements, but stop short either in the how many of the senses they use for data analytics, as for example in Newbury et al. (2021) or Yang et al. (2020b), or they create a convincing multi-sensory environment where the focus is not on providing analytical tools to interrogate the underlying data, as in Edler et al. (2019) and Hruby (2019b).

The reason for this lack of examples is of course that even the individual component parts of such systems are difficult to build on their own, and thus most research tackles those individual aspects first. There is a lack of specific research into cross-modal effects, there is a lack of off-the-shelf, ready-made hardware to display to multiple senses at once, there is a lack of software frameworks to help with this task, and a lack of techniques to display to multiple senses at once. The non-visual senses also bring with them distinct cultural and socio-demographic connotations, where different audiences are going to have radically different interpretations of certain sensory impression (Kühne and Edler, 2018). Despite these difficulties, working systems can be found in the literature and show some of the possibilities that such tools can unlock.

Perhaps the project that comes the closest is the "Tangible Landscape" project shown in, for example, Petrasova et al. (2018), or used for infectious forest disease forecasting in Gaydos et al. (2019) and terrain analysis in Millar et al. (2018). This system has been combined with immersive display hardware in Tabrizian et al. (2016), and allows tangible interactions with geospatial data through a special tracked surface, on which tangible elements can be placed or hand gestures can be performed to trigger analytical operations in a virtual environment that is connected to this table. Some impressions of this can be seen in Figure 2.6 and 2.7.

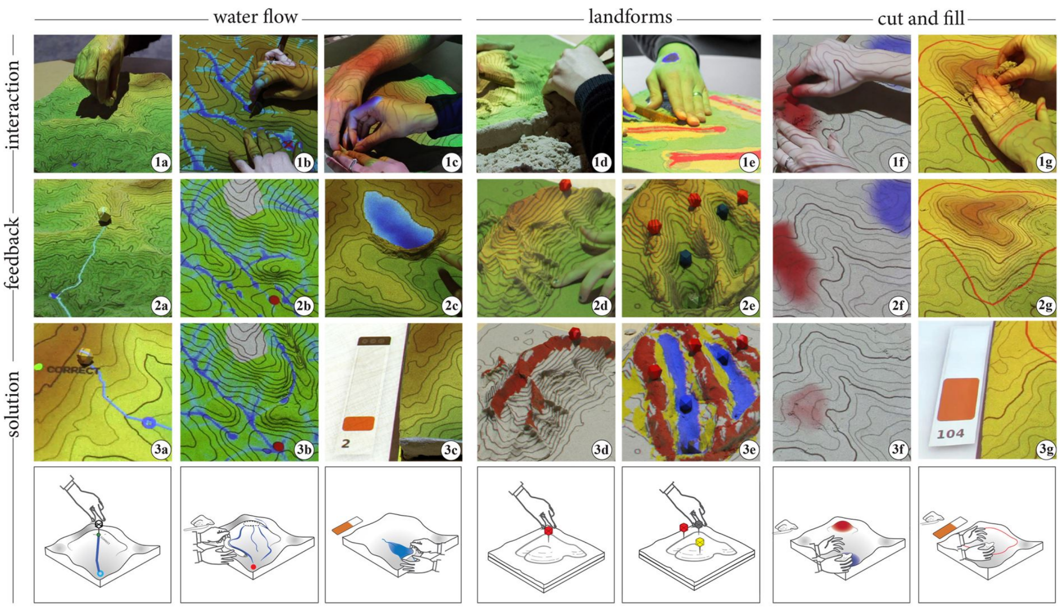


FIGURE 2.6: Embodied geospatial analytics on the Tangible Landscape system, shown in Millar et al. (2018).

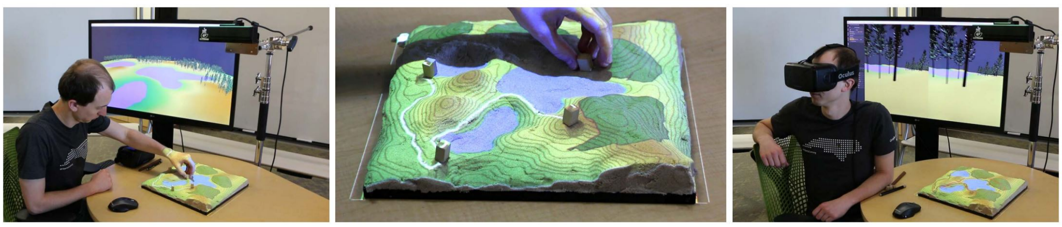


FIGURE 2.7: Switch to an immersive view in the Tangible Landscape project, shown in Tabrizian et al. (2016).

Work on another desk-based system can be traced from Harding et al. (2002) to Harding and Souleyrette (2009). Here, visualization, sonification and haptification are combined as a user draws on a map with a haptic stylus. The encoding of values in this map is very traditional: there are multiple layers, whose display mapping can be changed. The applications are equally traditional: in Harding and Souleyrette (2009) specifically, the system is aimed at road planning. The novel part is, that a multisensory feedback loop is introduced. Roads can be drawn by running the stylus over the map, during which the stylus' force feedback constrains the tip to the displayed map surface. Only one map layer is shown visually. The other layers can be mapped either to the haptic or to the sonic channel. Both haptification and sonification happen at the tip of the pen. The sensory mappings for haptification are: bump mapping, friction (over surface features), attraction (towards line features). The sonification was encoded in either pitch or complex sound arrangements that imitate the natural sounds of a land use class. Alternatively, audio alerts could be played if the pen tip crosses an existing road. Even though multiple channels are available for each of the three senses, every sense was only able to display one layer.

Finally, there are the case studies presented in Chapter 5, based on previous publications (Berger and Bill, 2019; Berger, 2020, 2021a). Only Berger and Bill (2019) contains a full embodied geosensifications as defined in this work, but both Berger (2020) and Berger (2021a) are conceptualized within the context of such systems and try to investigate sensification techniques or embodied interactions that would enable them.

2.4 Conclusion

In this chapter, we traced the technological and conceptual lines that lead us to the concept of embodied geosensifications. Section 2.1 highlighted how novel display and especially interaction technologies have created entire new sub-disciplines and research areas in the field of cartography. As a natural evolution of 3D geovisualization, which already had a heavy focus on new rendering strategies and interactivity, immersive geovisualization breaks with even more of the established strategies of cartographic representation. This section also first introduces the concept of cartographic variables, which will become very important at the beginning of the next chapter.

To be able to put these issues on a solid technological foundation, Section 2.2 establishes the technological space from which these new cartographic trends originally stem. It clarifies how many of the common terms will be utilized for the rest of this thesis and introduces the basic ideas behind important unusual technologies like non-visual displays. Specifically it introduces the concept of spatial computing technologies as a umbrella term for the relevant hardware and software solutions, and embodiment as an interface paradigm that this thesis will focus on.

Section 2.3 then took the topics established in these sections, showed the different ways in which they are already being brought together in the literature and where crucial research is still missing. It establishes many of the distinctions that will be important later on, like geographic scenes and their situatedness, as well as how certain concepts can be evolved, for example the move from visualization to sonification.

The friction that is generated at these intersection points is what the rest of this thesis builds on. A solid foundation in these topics was necessary to establish the context in which decisions are made throughout the next three chapters—beginning with the concept of using multiple senses for data display in immersive environments.

Chapter 3

Taxonomies of Embodiment and Spatial Data

When starting work on a project, it is often a good idea to assume a certain theoretical lens, in order to be able to make some assumptions and predictions about how the project should be approached. In computer-science related fields, we often default to technical conceptual frameworks like the Unified Modelling Language (UML) to guide our path. However, for something as complex as the embodied geosensification interfaces discussed in this thesis, these do not operate at the correct level of abstraction—before we can make decisions about implementation strategies, we first need to find the strategy behind the technique we are trying to implement.

In order to create new conceptual tools, this chapter will feature several discussions of different possible theoretical lenses we can use to think about different parts of the application we want to create. These lenses concern how we sense our surroundings, how we view space, and how we interact with these surroundings. This does not mean that such an application shouldn't be thoroughly modeled in software engineering frameworks, especially if it will grow complex, but it is just one conceptual framework of many that we should allow ourselves to use.

The lenses through which we look at embodied geovisualization correspond to the three aspects that every such system needs to cover: senses, data, and interactivity. These three will be looked at on their own in this chapter, and then be combined piece by piece in Chapter 4, in the hopes to attain a model that is both simple enough to be usable, but deep enough to spark new thoughts and experiments.

3.1 Senses—Sensory Variables and Sense Aspects

"There can probably be no complete agreement within cartography (or in information graphics as a whole) concerning what fundamental variables we have to work with. [...] In addition, the potential variables are dependent upon the technology available for producing sign vehicles and entire displays."

Alan M. MacEachren, How Maps Work (MacEachren, 1995)

At the heart of both modern visualization and cartography lie visual variables—features of an image or part of an image that are salient to our visual cognition in some way. Some of them catch our attention, some of them remain visible even in busy environments and others allow us to accurately resolve differences when a variable is used in multiple places or distributed over a surface. The development of this concept started in the purely visual domain of cartography, and as such that is where this section will start. Later, we will move from visual variables to sensory variables and, accordingly, from speaking about visualization to sensification.

While the concept of visual variables is often only cited as a list of seven variables with different attributes, in Bertin's original work itself it exists within a larger theoretical construct called the "Matrix Theory of Graphics" (Bertin, 2000), deeply enmeshed with considerations about semantics and semiotics. Bertin bases his description on an abstract description of a space with three axes X, Y, and Z. Within this space, one starts with a matrix of data (conceptually similar to the attribute tables that modern GIS is often based on), i.e. the semantics of the map, and then transforms this matrix into a graphical representation that is described as a 2D space (axes X and Y) with signs representing the data, i.e. the semiotics of the map. These signs placed in 2D space create the third axis (Z) through their "retinal variables", i.e. the nature of the light a viewer will perceive on their retina at this point in the graphic. Bertin spent decades further developing this approach, extending it and breaking it down into concrete steps that were tailored to the sometimes mechanical, sometimes digital means of map creation at the time (Bertin, 2000). His complete works did not gain much traction outside of the french cartography community (Harvey and Losang, 2019), but did in part inspire work on a grammar of graphics by Wilkinson (2005) and later Wickham (2010), that to this day remains the basis for a lot of applied data visualization, especially in the domain of data science. In cartography itself, Bertin's seven variables remain the most important contribution: the *position* on the two axes of the graphical 2D space, and the retinal variables *size, value, color, shape*, *orientation*, and *texture*.

Within these variables, he identifies four properties of such variables. The first of these is *selectivity*. When a selective variable is used the resulting symbols can immediately be perceived to fall into a group, even when all of them have additional other variables that separate them. *Associative* variables do not change the perceptual emphasis of an element and thus allows us to group elements without ordering one group over the other. Dissociative variables can be *ordered* or *quantitative*. Ordered variables can be perceptually ranked and quantitative variables can be used to perceive ratios between symbols.

Alan MacEachren builds upon this work in "How Maps Work" (MacEachren, 1995), which he describes as "taking a cognitive-semiotic to maps and mapping". He discounts the view of maps as communication devices that carry a predetermined set of messages and that can be measured for effectiveness and error rate like a communication channel would. He calls this approach "map engineering". Instead, he describes the act of making a map as an act of knowledge-making, in which a cartographer shows a part of the world in a specific way that is always colored by their subjective experience, artistic sensibilities, and personal goals, and the act of reading a map as a cognitive system perceptually processing a visual artifact and integrating it with their prior knowledge ("human-map interaction"). Maps are representations that carry meaning at many levels, most broadly at the public level (epistemological-philosophical-sociological-historical and logical-categorical) and from the private level (perceptual-cognitive). As a result, he starts collecting evidence from a variety of related disciplines, including psychology, sociology, semiotics, cognitive science, and more. Where Bertin constructed his theories based on experience, but still took a very formal approach to mapping, MacEachren tries to ground his broader hypotheses about maps in specific evidence wherever possible—which even ended up offering supporting evidence and extensions for some of Bertin's conjectures. With the visual variables specifically, MacEachren arrives at the following list: *size, shape, value, orientation, texture, location, hue, saturation, arrangement, focus, resolution*, and *transparency*. Hue and saturation are simply Bertin's color variable split in two, and the three variables arrangement, focus, and resolution are added. The variable of focus specifically is intended to represent data uncertainties, something that Bertin did not deal with in his work.

MacEachren's representational view of maps actually carries us further towards the kind of cognitive considerations that we have already seen in HCI and spatial computing research. Especially the idea of sociological factors influencing map reading also immediately raises questions about using non-visual senses, because as noted earlier, many of them are considered much less "objective" than the visual. In some ways the auditory or the olfactory operate on the sociological factor as much as they operate on the perceptual. How, for example, would people of different cultures interpret the smells that are assigned to different areas in a hypothetical "odor map"? Every single smell might trigger different memories, have religious connotations, be a reference to different hallmark dishes, and thus be received in vastly different ways.

Despite all this work, many hypotheses specific to map signs remain to this day untested in an empirical way. There is some evidence for the perceptual effectiveness and ordering of individual variables, especially hue, size, and orientation, which lines up with Bertin's conjectures in an almost surprising way, however most studies in this area assume a a very simple rendering scenario instead of complex maps with multivariate attributes (Garlandini and Fabrikant, 2009). Many common cartographic wisdoms are instead based on experience and different semiotic theories. The evidence gets even more tenuous as we move to the non-visual sensorities.

Finally, we could also debate the usefulness of anything so simple as *variables* for visualization. Sometimes they seem like an artifact of a bygone era, as visualizations grow ever more complex and interactive. Often the data mappings are not simple enough to be compressed to such a concept, especially in immersive and situated applications. The reason why they still hold at least in the geospatial case, is that the features in the geospatial domain are there, already located and have a shape. As such we already have a clear reference within our immersive geographic scene, the same way we have clear X and Y coordinates in 2D cartography. These features might change their shape and location in the scene over time, but they must always be identifiably present, in order for the user to be able to analyze them in their geographic context. From this clear reference in space, sensory signals can then originate. This makes the value mapping process more constrained than in abstract data representations.

It is important to note, that these variables describe how we use perceptual qualities to encode specific information. That does not mean that the underlying perceptual qualities will *only* be used for such variables. In fact, we need to be careful to consider overlaps between different uses. The color green could be used to communicate some attribute value in the data, but in a virtual environment it could also be the color of the grass on the ground, which might have no relation to the data attribute whatsoever.

3.1.1 Variables in 3D Cartography

MacEachren builds much of his model of cartography on semiotics. This basic relation remains, even if we move to 3D or immersive environments. Instead of presenting a viewer with a system of signs on a piece of paper or display, we put them into a sort of 3D "sign world".

Carpendale (2003) is perhaps one of the earliest publications that makes note of possible new variables enabled by 3D space: *depth, occlusion, aerial perspective, binocular disparity, stereo viewing*. However, several of these variables specifically refer to visual effects as created by a conventional 2D display. To dynamically alter depth perception, binocular disparity, or the parameters of a stereo display while wearing a visually immersive display system could quickly induce motion sickness and eye strain in users. We need to be careful to not lose track of our focus on immersive technologies—in an interactive 3D environment variables like depth and occlusion are just an effect of object and user position, instead of a channel that data can be displayed over. There are however possible edge cases in which display depth could be altered to diverge from object position, for example to highlight a specific element in a 3D point cloud. In order to cover both depth and occlusion in a way that corresponds with the actual working of computer graphical systems we will introduce the variable *culling*, which includes every divergence from the depth culling that is expected in interactive 3D environments.

Apart from this effect, the changes between 2D cartography and 3D cartography are simply not relevant enough for immersive geovisualization—the minutia of this reprojection effort do not make for a good selection of variables. Instead, the immersive, multisensory variable space needs to be considered on its own merits. Much of what might qualify as a 3D variable, is instead a function of the variables in the next subsection: location and time.

3.1.2 Location & Time Variables

We have to discern two types of location in spatiotemporal visualization: The location of an object itself, which is given by its movement dynamics as shown in Section 3.2, and thus not a part of this chapter, and the sensory variable of location—the specific display location of a value signifier on or near the position of an object. This latter concept is highly dependent on which sense is noticing an object, and will thus be considered as a sense-specific variable in Section 3.1.4.

A similar distinction holds for the concept of time. An object might move through space according to its actual change over time, but for visualization purposes it could be made to stop (stutter) or disappear (flicker) to signify some value change. This is a lot less dependent on what sense is noticing the change over time, thus this section will consider the temporal variables in their totality. At first there will be a quick introduction of human chronoception, as it is in some way the root of all human sensory cognition and has a big impact on the other variables.

There is no centralized sensory system for feeling time, instead there are distributed internal systems for different kinds of timekeeping (Rao et al., 2001). None of them are as precise or reliable as we often expect our other senses to be, as they seem very contingent on the other contents of cognition over a measured time span. Time may seem to run faster when engaging in a stimulating activity, and even alterations in the light of a room can delay the circadian clock. Knowing this, it is not surprising that time often feels highly subjective. There is much philosophical and scientific debate about the nature of time, how it comes that we are pulled along it as we are, if it has a direction at all, or whether causality is an artefact of our cognition more so than a fact of the physical universe.

Fortunately, solving the true nature of time is not necessary for useful geovisualization, not even in the immersive case. Instead, we can start at a very different layer of abstraction: DiBiase et al. (1992) name three dynamic (time-based) modes of cartographic expression: Sonification, Interaction and Animation. Sonification will be explained in Subsection 3.1.4.2 and Interaction in Section 3.3. Here, the focus shall be animation, which will be taken to describe changes to the sensory display of an object, that will usually happen over clearly perceivable time frames from fractions of a second to multiple seconds, as opposed to the often much faster alterations in a sonic signal (more on this in Section 3.1.4.2). For the purposes of generalization we can call this aspect of sensification its *dynamics*. A sensification can for example be *interactive* without including any *dynamics*, if no object ever moves or changes shape to represent some data state, and instead remains static up until the point where it is directly moved by user interaction.

A dynamic sensification fundamentally is a sequence of static sensifications. The three variables relevant to such a sequence according to DiBiase et al. (1992) are *duration*, *rate of change*, and *order*. *Duration* refers to how long one static impression in the sequence is shown. A change in duration can have an impact on the pace of a sensification. *Rate of change* refers to the degree to which all the other sensory variables change over a certain duration. *Order* describes how the sequence of static impressions is arranged—with spatiotemporal data, this will usually be chronological, but it does not have to be. In MacEachren (1995), *display date*, *frequency*, and *synchronization* are added. *Display date* refers to when in an animation the sensory representation of a feature is shown, *frequency* describes the number of sensification states shown over a certain time frame (for example, 3 state changes per second), and *synchronization* describes how state changes over multiple features correlate with each other.

Köbben and Yaman (1995) also runs some limited tests on the perceptual properties of these variables. Notable is disagreement in some of their property assignments as opposed to MacEachren (1995). They for example list display date (called "moment" in this paper) as non-ordinal, while MacEachren (1995) assigns it a possible effectiveness as an ordinal variable.

These time variables work well for most maps and reflects the way in which animation is usually handled in GIS software and even real-time 3D engines. There are however some intricacies of immersive worlds which might make this too reductive. For example, most immersive environments will include some degree of simulation, like physics simulations or other, more semantic object and scene behaviours. These simulations can be non-deterministic, non-interactive, and often depend on a large number of parameters. As simulations they are always tied to the passage of time. To the time variables as proposed by DiBiase et al. (1992) and MacEachren (1995) we thus add the variable of *agent behaviours* (including both the laws of nature and the dynamic behaviour of objects). Using behaviours as a data variable instead of as presence-inducing background element could for example mean to let artificial agents like animals act more frightened the closer a forest fire approaches. Paired with a variable that describes the physical simulation properties of objects and nature, like the amount of gravity over a landscape, these simulation variables are what would drive ecological sensifications as shown in Moloney et al. (2018). These simulation properties however will be a part of haptification in Subsection 3.1.4.3, as they are also relevant without time, for example when picking up a virtual object.

While we distinguished these variables from the kinds of timed changes that create a sonification signal, every one of the aforementioned time variables is obviously also relevant to auditory signals, specifically once we move from pure sonification signals to data display through musical compositions. An addition we thus need to make, that is not traditionally made in timed visual signals, is the variable of *quantization*. This refers to how continuous timings in other time variables like frequency or synchronization are turned into discrete timings according to a (usually scene-wide) beat. The reverse way to state this would be how "off-beat" a timed signal is allowed to be. This reference beat allows for a more hierarchical value encoding as compared to synchronization, which would be more relevant for grouping similar variables.

Another variable that is at home in audio, but is more about the high-level arrangement of certain other variables over time is the concept of *composition*. In sound, this could include different scales, sequentiality, melodies, melodic leads, articulation, harmony, chord progressions, spectral duration and reverb effects. Dubus and Bresin list an exhaustive list of possible factors in Dubus and Bresin (2013). In the visual domain, this could describe a more complex play of colors and shapes, as one might find in a music visualizer software.

This completes the temporal aspects of sensory mapping. For the rest of this section, we will establish a clear enumeration scheme for all our variables, numbering them in relation to the sense they operate on. The time variable are thus:

T1: Duration

T2: Rate of change

T3: Order

T4: Display date

T5: Frequency

T6: Synchronization

T7: Agent behaviour

T8: Quantization

T9: Composition

3.1.3 Linguistic Variables

Often not part of considerations about visual variables or sonifications, but a big part of them as actually implemented, are direct language-based outputs like speech, text and braille. Because our understanding and cognition of the world is heavily entangled with language use, basically any meaningful fact can in some way be presented in language. Because this is often used in combination with other variables (like a text that has a certain location or color) and because immersive visualization does not happen in relation to some form of readable "document" like non-immersive sensification usually does (like a printed map or an embedded map on a website), we will consider this a variable like any other for immersive environments—in fact, there are two: **speech (l1)** and **writing (L2)**.

3.1.4 Sensory Variables

Long after MacEachren et al.'s work on visual, haptic and audio map variables, and long before Marriott et al. (2018) seminal work on multisensory analytics, Jonathan C. Roberts and Rick Walker called for a "Unified Theoretical Approach to Multi-sensory Information Visualization" during a Workshop in the 2010 IEEE VisWeek (Roberts and Walker, 2010). In this workshop, a holistic space of perceptual variables after Bertin's visual variables was proclaimed as one of the challenges, amongst questions about data enhancement for multiple sensory modalities as well as cross-modal interference. Jonathan C. Roberts went on to be one of the main authors of the Multisensory Immersive Analytics chapter in Marriott et al. (2018), which incorporates many of the aspects that were called for in this workshop. However, a holistic space of variables is not included, perhaps because of the quite complex and multi-layered nature of immersive visualizations. As argued in the introduction to this section however, geospatial visualizations constrain the problem space enough so that the concept of variables becomes more relevant, similarly to how the variables were created in the discipline of cartography and not visualization as a whole.

Thus, this section will be a solution proposal for what Roberts and Walker originally called for: we want to collect all sensory variables that are relevant throughout all the primary and potentially even secondary senses. However, we constrain this specifically for the use case of spatial computing applications, i.e. applications with a located user and external data objects on which values are displayed. As imagined in the original figure (see Figure 3.1), we will spend some time considering display device technologies as well as design aspects and transference from visualization, however the focus will be on the holistic variable space. Sensory integration between multiple senses will not be the focus, as research is still severely lacking, but will be included at the end of this section. The whole section will follow the order of senses in Figure 3.1 and also use the five sensification sub-disciplines as identified there for section titles and variable names: V for visualization, S for sonification, H for haptification, O for olfaction, G for gustation.

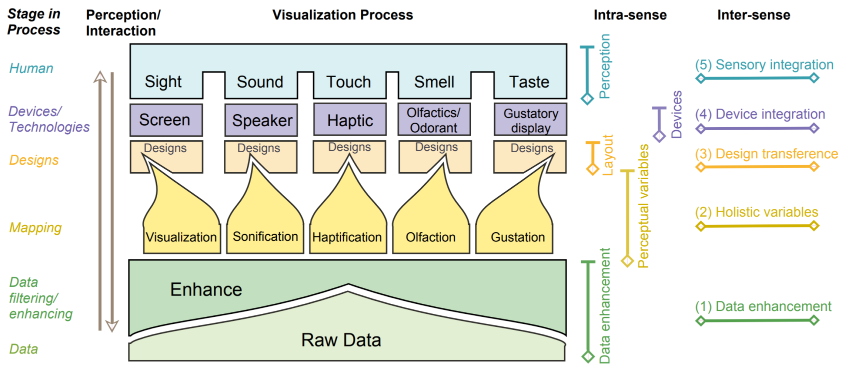


FIGURE 3.1: Overview of different facets of multisensory data visualization, shown in Roberts and Walker (2010).

Together with the sensory variables we will also list and discuss the spatial variables, which are a lot closer to how the individual senses operate than temporal and linguistic variables (more on this in Section 4.2.7), but still are a different class of variable in the geospatial case. They will be explained for each sense separately, but we will list them here to establish their numbering. Because the letter S will be used to denote sonification variables, we instead use Sp:

Sp1: Location

Sp2: Size

Sp3: Orientation

Sp4: Shape

Sp5: Environment

3.1.4.1 Visualization

For most people, the visual sense is their main sensory interface with the world. At its most basic level, it is comprised of two front-facing sensory organs sensitive to light, which both are capable of perceiving an elliptical 2D image. That image is very sharp in one area near the center (the fovea) and degrades in resolution towards the outside. Only the inner part of this ellipsis carries color information. Both of the sensory organs (eyes) are connected to the nervous system by the optical nerve, which creates a blind spot within the field vision. From the optical nerve to the responsible brain centers, lots of neural processing is done to enhance the two images and merge them into an image with depth information.The depth information is based on monocular depth cues like relative size, shadows, motion parallax, and stereo depth cues like binocular parallax (based on differences between the two 2D images) and convergence (based on the state of the oculomotor muscles used for focusing on objects).

Hardware

The necessary technical foundations for these displays have already been covered in Section 2.2. However, one distinct issue needs to be highlighted here: Not every visual display will be able to deliver every visual variable to a viewer. Different display types feature different qualities of colors and resolution, and not all displays in immersive environments necessarily can display depth. An optical see-through display might have issues with displaying proper brightness levels in broad daylight, thus the brightness value can not be used as a display variable. Similarly, any sort of projector will be unable to display black, unless the medium it is projecting on can be made black itself. Any sort of ranking of these variables always either has to take into account context or assume perfect display conditions. In many cases, creating these conditions is a balancing act between cost, ergonomics, and power supply/heat.

Variables

Visual variables were already discussed at the beginning of this chapter, specifically those established by Bertin (1967) and MacEachren (1995). What follows is our specific selection, based for the most part on the variables as defined by MacEachren, but adapted to immersive environments:

Sp1: Location The place at which an object is displayed. This will usually coincide with real-world location of the visualized phenomenon—however, when used as variable to display some attribute value, one or multiple axes of location can be altered to instead encode information. Often this can mean replacing the height values, if we have a ground-level data set. If there is some sort of regular pattern or raster in the placement of the objects that can serve as a reference, small irregularities in the location could also be used for encoding. In combination with time variables this variable also describes motion.

Sp2: Size Changes in the size of an object. As in cartography we have to be careful about how this change is mapped—doubling the radius of a sphere creates a disproportional increase in volume, while doubling the volume will not necessarily appear to the user as a doubling in the mapped attribute. In combination with time variables this variable also describes scaling.

Sp3: Orientation This encodes the direction in which an object is pointing in 3D space. This is a visual variable in the same sense that location is—it determines what parts of an object are *visible*. In combination with time variables this variable also covers the concept of rotation.

Sp4: Shape This variable can refer to much more complex arrangement than the previous spatial variables. It can also be semantic (for example when objects take well-known shapes representing some fact about the data), but it does not have to be—a shape in 3D space can be any manifold. In order to be readable as a data variable, the possible space of shapes has to be limited and well described to a user. Possible variables that can be derived from shape are mostly topological facts like “roundness”, “edge sharpness”, “number of holes”, and many more. In combination with time variables this variable also covers the concept of shape-changing.

Sp5: Environment This newly introduced variable covers any effects that describe how the visual display of an object interacts with and alters the appearance of the environment. This is relatively rare in visualization, but can describe factors like data display through ambient light or alterations to light propagation behaviour in virtual environments.

V1: Color value Also called "brightness", this value encodes values to a spectrum reaching from black to the brightest version of whatever hue and saturation the object currently has.

V2: Color saturation This value encodes to a spectrum from white towards the most intense version of the current color configuration.

V3: Color hue Here, we encode the “color”, i.e. not a relation to black and white, but the actual tone of the color. This is the outer edge of a color circle flattened onto a line, i.e. a spectrum moving from one hue, through all the other hues and back to the beginning.

V4: Texture Especially important in 3D worlds, this describes any way in which procedural generation of surface materials for 3D objects can be used to encode data. These procedural surfaces usually have many parameters, some of them are other visual variables. Even though this is in this way often a “derived” variable, we still describe it here because the encoding of data on textures can go far beyond changes in colors or noise patterns—textures bring with them semantic possibilities, like encoding a data fact in how “metallic”, “reflective” or "bumpy" a surface looks, even when these semantic descriptions are ultimately just a complex arrangement of other visual variables. Here, the implicit semantics that we bring with us from a real world context are the point more than specific quantifiable attributes.

V5: Blur This describes any process that impedes clarity of vision on an object or any random destruction of data quality, where the purpose of this impediment is to highlight some fact about the data, usually some form of uncertainty within it.

V6: Transparency A historically difficult part of computer graphics, transparency makes it possible to see objects or scenes behind an object. In visualizations this transparency is sometimes used to circumvent self-occlusion of data sets. If the data set does not have occlusion or the occlusion is handled in some other way, the transparency can instead be used to encode some data fact.

V7: Level of detail In comparison, here we have a very deliberate and structured lowering of data quality or resolution. In visual encodings this is usually done for performance reasons or to arrive at a specific art style, but if used as a data variable it could be used to rank data objects by a value and make comparisons between them easier.

V8: Culling Occlusion is a natural fact of 3D scenes. However, some objects may have visualization priority and thus could be rendered even if they would normally be occluded by another object—virtual scenes can and do break the laws of reality in this way. This priority could possibly be used to visualize some fact about the data, most likely through some interactive process, like a virtual lens that a user could hold over an object to see through it. Lets imagine a dataset with a large amount of data points represented by spheres. If the spheres are too close to each other, a culling visualization could allow us to use a see-through lens to make those spheres visible that fulfill a certain filtering condition, without having to make all the other spheres transparent.

Building on Bertin (1967), Morrison (1974) splits color into hue and saturation, which is the distinction adopted here. In practice, the three color variables are however often reconfigured and combined into palettes, i.e. continuous series of certain changes in all three factors, in order to encode some data fact, instead of being used for direct encodings. The reason for this is that our color perception is so rich with differences and strange sensory adaptions, that the direct mapping will often feel wrong or be difficult to read for most people. Specifically selected, limited palettes can follow aesthetic concerns, but they can also be there to make changes in a variable even readable in the first place, for example for colorblind people.

It is likely that there are complex interrelations in all these visual (and non-visual) variables that we just do not know yet, because only color has had decades of focused research spent on it.

Blur and level of detail are the renamed versions of crispness and resolution within MacEachren (1995). They were renamed here to bring them more in line for parallel use with the other senses or to make them more applicable to 3D space and 3D rendering.

3.1.4.2 Sonification

Where sight is limited to the front of our body, hearing resolves vibrations reaching us from every possible direction. Just like sight however, it is a sense that works on the stereo-principle of having two sensory organs of a fixed distance to each other. A 3D impression is then generated through two steps. First, a 2D stereo position is derived from interaural (inter-ear) time difference (ITD) and interaural intensity difference (IID), and can be thought of as spanning a 2D plane through our ears and orthogonal to the axis of our body, to which all sounds from 3D space are then projected. In order to resolve an actual 3D position, we use two more cues: monaural cues derived from the way that sounds of certain frequencies and directions interface with the shape of or ear and the ear canal, and dynamic cues that happen as head movements shift the "stereo-plane" through space.

Hardware

There are many forms of audio display systems, from speaker systems to headphones and ear buds. They all bring their own challenges, but they all have one main, largely unresolved, problem within them: It is easy to display a sound on the 2D stereo-plane and difficult to display a 3D position. In order to accurately project 3D sounds, a speaker system needs to be turned into a full omnidirectional speaker array. Headphones and ear buds on the other hand, even ones with multiple speakers in each shell, are often too close to the ear to fully imitate omnidirectional sounds, and can not create the full head and upper body resonances one experiences within a speaker system. For this reason, to display 3D sounds accurately, sounds need to be run through a head related transfer function (HRTF) that can be tailored to individual bodies to pre-modulate the sound in a way that when it reaches the inner ear, it appears correctly modulated by the outer ear and surrounding bones and tissue.

Variables

Audio-maps appear early in the literature and are even mentioned in MacEachren's seminal work (MacEachren, 1995). The now relatively common term of "Sonification" was not in use at the time, but the variables that MacEachren focuses on, taken from Krygier (1994) are: *location, loudness, pitch, register, timbre, duration, rate of change, order, attack/decay*. The occurrence of the time variables in these strikes directly at the core of one of the specificities of auditory signals: They are always already temporal. One might rightly argue that all cognition, and thus all senses, need some time to process the signals that reach them, however what makes sounds special is that it is incapable of detection something that is not dynamic—the change of the signal *is* the sound. A non-dynamic object is just inert air pressure—complete silence. What instead makes a sound signal, are the micro-scale dynamics happening over short, sometimes even imperceptible, amounts of time. A common guidelines here is the range from 20 Hz to 20,000 Hz. Anything lower than this range can be seen as composition of individual sounds—which is where we switch from sound design to musical arrangement. Much of sonification attempts to have musical characteristics, in the same way that most maps also try to be visually pleasing. However, in these macro-level cases, it is more useful to instead use the already established time variables like composition and synchronisation.

A much more detailed version of the audio variables can be found under the title of "auditory dimensions" in a review paper by Dubus and Bresin (2013). They identify 30 variables in 5 categories. The categories are: pitch-related, timbral, loudness-related, spatial and temporal. All of these categories are direct matches to variables we already identified. The individual dimensions however are way too specific for our purpose, for example including technical details like the distinction between sound spatiality through HRTFs, stereo panning or vector base amplitude panning. More interesting are the distinction in the temporal variables, which include concepts that have not been considered before, like different kinds of duration: "Ambient duration" for long-lasting, slow changing signals, or rythmic duration for durations between 100 ms and 2 s.

Ultimately however, many of these dimensions are of a very technical or musical nature and too specific for our more abstract taxonomy. More useful here is a distinction made explicit by Edler et al. (2019): Mapping data to sound can be realized through voice, through music, or through abstract sounds. Voice is covered by the language variable. Thus the question is: what do we need to conceptually capture both music and abstract sounds? Or rather: What is not covered by just using Dubus and Bresin's five categories directly? Our spatial and temporal variables have already been established. Pitch and loudness can be used directly. Timbre is at a similar level of complexity to the visual variable *texture* and thus should not be broken up further. Spatiality has a very different connotation in our geospatial system and thus should not be looked at from the domain of stereo rendering. Most immersive systems are going to default to HRTFs and just accurately display the position of a data point. Which leaves the temporal dimensions.

Again, we should look to texture and timbre: anything that is so domain-specific and complex that it is a combination of other variables but is too complex to be easily described by them, should be rolled into its own variable. However, if a variable gets complex enough that it starts being about the interrelations of other variables over time, it should instead be part of the temporal variable composition. The only larger-scale audio variables that would not fall under compositions are variables that could affect an individual tone too, i.e. are not temporal on a scale spanning multiple tones and also do not fall under pitch, loudness or timbre. This arguably includes *Reverb*, which is a temporal and spatial effect that can affect an individual tone and could very easily be used to encode some data fact. Reverb, and related effects like echoing, however are sufficiently covered by the spatial variable of environment, as explained below.

It also includes a fully temporal dimension only partly mentioned in Dubus and Bresin (2013): the signal envelope, usually called the ADSR (Attack, Decay, Sustain, Release) envelope. It appears in Krygier (1994) as the variable *attack/decay*. This is much more small-scale than aspects like melodies and resolves to our ear more as part of an individual element in a musical composition. If we were to, for example, encode a value as a pitch, different envelopes would make that pitched tone appear very differently to a listener.

Interestingly, Dubus and Bresin (2013) leave out one aspect of audio signals that is extremely important in musical arrangements: clarity. We often add *noise* to signals or *compress* their higher frequencies for "warmer", less sharp sounds. These clarity variables roughly correspond to visual variables like blur and level of detail and could be treated as analogues for sensification purposes.

Thus the final selection of sonification variables in immersive geosensification is:

Sp1: Location Just like with visual perception, auditory objects can be moved off of their real-world position to encode some information. This movement can not be as subtle as in vision, as locational accuracy in hearing is less precise than in seeing.

Sp5: Environment Environmental effects are very common in audio processing, including reverbs and echoes generated by virtualized rooms. In virtual environments this will usually be a physicalized feature used for increasing immersion and presence, but it is also very common to artificially alter such factors, most commonly in "synthetic" reverberations—and as such it can be used for data sonification. Almost no modern piece of music does not heavily feature reverberation on its individual sounds, so listeners are trained on this kind of manipulation. One way this variable could be utilized is by imagining a virtual room around a sound source and then simulating reflections and diffractions in this room. These artificial effects can then be rendered into the sound signal the object emits. This sound could of course reverberate further as it travels through the environment, but if applied correctly it is relatively easy for us to distinguish what is an artificial sound feature and what is a physical effect.

S1: Volume The simplest modulation of an audio signal is over its energy. Volume changes are easy to recognize for everyone, however the variable is constrained by the range of what is safe to listen to and what users with various hearing impairments are capable of resolving.

S2: Pitch Often shown to be the most potent auditory variable for encoding the up and down of data like a conventional cartographic visual color scheme would, pitch refers to whether a tone is high or low. When generated by a musical instrument, a pitch is defined by the fundamental frequency, i.e. the lowest partial. The rest of the harmonics in the signal, if they are present, are part of the timbre variable. Pitch is based on frequency, but differs from the frequency variable in that human psychoacoustics permit us to perceive the frequency of sound waves as an immediate tone instead of a change over time. A reoccurring change in pitch over time such as generated by LFOs (Low Frequency Oscillators) in sound synthesizers is thus defined by the frequency (T5) of a change in frequency (S2).

S3: Timbre Like the texture variable in the visual domain, this variable offers a plethora of possible signal modulations. While there is no real way to order values by this variable, apart from subjective measures of pleasantness, it is very potent in distinguishing different configurations. We have all heard the spectrum of sounds that even a single instrument can create—volume and pitch do not even begin to describe the totality of its possible outputs, the same way that hue and brightness do not describe the appearance most real world surfaces. Timbre is in part a side effect of all the other audio variables and their combinations, but specifically is defined by the full spectral envelope of a signal, i.e. not just the dominant pitch, but all the layers of lower and upper harmonics that modulate this envelope. As far as technical precision makes sense with a highly qualitative variable like this one, a clean sine wave could be said to have no timbre, or less controversially a very simple timbre, while the sound signal of a full orchestra has a highly complex and layered timbre. Dubus and Bresin (2013) specifically includes in this variable the subvariables of instrumentation, polyphonic content, voice gender, allophone, spectral power, amplitude of harmonic, frequency of harmonic, roughness, brightness, center frequency of filter, and saliency.

S4: Envelope The combined aspects of attack, decay, sustain and release. This describes how fast an individual tone reaches its maximum loudness, how fast it falls back off and how long it takes to fully disappear.

S5: Noise Audio is where the term noise in its signal and sensory processing context originates. Noise is very important in audio processing, either as something to be removed, or as something to be added in controlled amounts. This is often done to make signals sounds more pleasurable, in case of a small amount of noise, or very raw and energetic, in the case of large spikes of noise. It is thus a potent variable for encoding information, as it is one of the less subtle aspects of an audio signal, especially when listening for changes over time.

S6: Compression Compression is quite common in audio production. It lowers the sound level at the higher intensities of an audio signal, while leaving the lower levels unaffected. Modern music specifically is often passed through multiple layers of compression to constrain the signal. This has the effect of producing a less sharp, often more pleasurable, output, as small spikes in signal strength are smoothed over. It differs from timbre in that the amount of compression is much easier to quantify and is usually applied as a distinct processing step. It is however often quite subtle and can be difficult to hear for the untrained ear.

Some additional notes on sound: Timbre and Pitch are often related in the human voice and physically existing instruments. Changes to one often modulate the other in subtle and not-so-subtle ways. This interlinking between pitch and timbre is usually called *Register* and is included as a variable in Krygier (1994). MacEachren (1995) includes it and notes that it could be useful for representing nested hierarchies of pitch values. In our categorization, register used in this way would however be a part of the composition variable, as it is usually a musical aspect rather than a signal property.

3.1.4.3 Haptification

Unlike the visual and the auditory sense, what is commonly referred to as the haptic sense is a combination of a multitude of senses throughout the entire human body. The involved parts of the body are part of the somatosensory system. This includes mostly exteroceptive senses, but also some interoceptive senses. Exteroceptive are the sense of touch (mechanoreception, also referred to as tactile perception), temperature perception (thermoception) and pain receptors (nociception). Interoceptive senses includes the sense of body position (proprioception, sometimes also referred to as kinesthetics, especially when considering motion) and internal pain receptors (Schneider and Feussner, 2017). Obviously internal pain receptors and body position are often affected by outside impulses (for example when our movements encounter resistance), which is why they are generally seen as part of the normally exteroceptive sense of touch.

For our purposes, these basic facts will be detailed enough, however it should be noted that there are many additional complexities to these senses. The different receptors described above can be located in many different body areas from the different skin layers, to hair follicles, the tongue, and some mucous membranes inside the body. Depending on where they are located their "calibration" can be different, and the same input force can thus be perceived very differently. If we wanted to go to this level of precision, we would have to differentiate between many different sensations, like for example: light touch, deep pressure, vibration, low frequency vibration, tickle, itching, hair movement, skin stretch, and more (Chouvardas et al., 2008). In the more abstracted view we try to adopt here, some of these will still reappear when considering the haptic variables.

Hardware

The haptic senses are perhaps the most unwieldy to create display hardware for, in the sense that we often need to create the display system with the application already in mind. The only generalized displays are those that can be worn and then try to simulate external objects through application of force to the right points on the body. For tactile (and possible pain and heat) sensations we have items of clothing incorporating arrays of vibration motors or heat cells, or displays that project pressure through air flow or ultrasound. For proprioception we can wear exoskeleton systems, which are capable of putting up resistance against the movement of our tendons, ligaments and joints. This differentiation into two kinds of haptic displays is very commonly adopted in the field and makes classifying these systems more manageable (Marriott et al., 2018). The distinction can however break down with tangible interfaces, which are physical objects located in the environment, and thus cause both tacticle and force-feedback sensations. For this reason, they are usually treated as a different class of systems compared to worn display systems. Marriott et al. (2018) solves this by distinguishing two approaches: data physicalization (data is represented by a physical object) and haptification (data is explored with devices that synthesize haptic variables). Modern technology however could also combine the two, with highly dynamic specially-made objects that feature actuators and motors.

If we have the chance to develop a haptic display specifically for one application, the possibilities increase considerably. For the case of tactile displays, Chouvardas et al. (2008) for example describe both mechanical energy devices, which cover technologies based on vibration, ultrasound, surface acoustics and even so-called electrorheological fluids (fluids that change in viscosity based on an electric field), as well as electro-tactile stimulating devices.

A big advantage of haptics is that most of our human-computer interaction systems are ultimately about registering limb- and joint-movements. Haptic displays can thus often double as input systems, as is the case with modern VR motion controllers. This advantage even holds for non-generalized haptic displays like tangible objects, as the space of interactions with the tangible objects can usually be made rich enough to not require any additional abstract inputs.

This can go even further: In the same way that haptics are often entangled with interactions in immersive systems, haptics are also deeply involved with out real-world use of tools. Once a human has become familiar enough with a tool, they can have some limited form of haptic perception *within* the tool, outside of the body. This sensory capability is referred to as *extended physiological proprioception* (Simpson et al., 1974). Because physical forces travel through materials, the feedback of our tool still invokes haptic signals within our body, which our cognition is capable of resolving as if they had happened within the tool itself. In immersive systems, this could of course be exploited: if we controlled the haptic outputs to the body closely enough, we could possibly attain this extended physiological proprioception within tools of arbitrary range and form—even with physically impossible virtual tools.

Variables

This is where we have to most substantially deviate the most from what was originally laid out in MacEachren (1995). Here, the haptic variables presented by Vasconcellos (1991), which were very much focused on tactile maps for the visually impaired and correspond almost perfectly with Bertin's variables, are front and center: *volume, size, value, texture/grain, form, orientation, elevation*. While some of these remain important, haptics suddenly grows much more complex and capable in immersive environments, but is also far less specific. A variable like elevation for example does not mean much if objects are complex and three-dimensional. We also have dynamic haptics like vibrations or even force feedback (mechanical resistance against movement) that introduce new, often temporal, variable possibilities into the space. These variables are often akin to some of the auditory variables.

Once we introduce these possibilities, Vasconcellos (1991)'s whole variable space collapses into a variable that describes the *surface* structure of an object, as the space of possibilities just grows too large and dynamic. For certain applications we could selectively use these variables to further describe the nature of one surface structure in the larger immersive space.

For a source with variables much closer to our applications, we can consider Griffin (2001), which specifically criticises and builds upon Vasconcellos (1991) in the context of virtual reality technology in the early 2000's. Griffin's haptic variable syntax includes variables that are touch-only (*vibration, flutter, pressure and temperature*), touch variables with direct visual analog (*location, size, texture/grain, shape/form, orientation and elevation*) and kinesthetic variables (*friction, resistance, and kinesthetic location*). These are however only examples and the list is never fully concluded—neither by Griffin nor, to the knowledge of the author of this thesis, anywhere else.

To finish this list, let us start with what can be discarded or reduced. Location, size, orientation and shape are covered by our spatial variables. Texture is similar to texture (V4) in vision and thus needs to be included. However, as we will see later, there is a better way to conceptualize this. Pressure, as in air pressure, is another variable that stays relevant. As already stated before, elevation loses its meaning and is included partly in texture and partly in shape.

Flutter is a specific modality that the surface of the human skin can detect, just like texture, pressure, different levels of vibration and stroking, as well as stretch (Chouvardas et al., 2008). These are however not variables that describe display properties of an object, but rather properties of skin. Our variables in this case should describe how the former can communicate facts to the latter, and as such all of these variables fall under the variables of texture or vibrations, and under temporal changes in either of them.

Finally, kinesthetic location is an intriguing variable, as it is insufficiently explained in Griffin's original paper. The configuration of the user's body itself could conceivably be used for data display, but only through active application of forces. The usefulness of this is questionable. More interesting is the opposite configuration. As the user's body moves, the kinesthetic position as represented in the virtual space can start diverting from the real. There needs to be a lot of care put into making this a comfortable experience, but the sensification possibilities are intriguing: such kinesthetic differences or "movement lag" could show faulty data configurations, facts about time or data quality. This data encoding through lag is not possible in other senses, as we usually do not have the ground truth that we have here—our proprioceptive feeling of body configuration can not be fully overridden by the virtual world.

From there, we can start with Griffin's syntax to find new variables. We start at touch-based sensations. Griffin specifically excludes the use of *pain* as a data variable. This is arguably a good idea, as there are potential ethical problems with causing pain for the purposes of data visualization. One use case however that is not within Griffin's area of application but would fit into a situated analytics context is safety-related haptification in extreme environments. Pain has the distinct advantage of being able to trigger immediate protective measures within the human body (moving away, covering vital areas). These are often pre-attentive, i.e. far faster than our consciously taken reactions. This could be useful in dangerous environments with normally painless but mortal dangers like toxic gases and radiation. We could for example imagine a geiger counter that starts to move from sonification to "nocification" once a body part moves into potentially lethal concentrations of radiation, causing an immediate flight response instead of auditory overload.

Then, there is the idea that vibration signals have direct overlap with audio signals—strong enough changes in air pressure often even trigger both senses at once. From sonification we can derive haptic analogues for volume (*intensity*), pitch (*frequency*), envelope (*abruptness* and *sharpness* of vibration impulse). Aspects like compression, timbre, and reverb however are too detailed to resolve for the haptic sense.

One entirely new variable can be derived from current work in haptic displays: *weight distribution*. While changing mass on the fly is often not readily available as a hardware capability, there are ways to shift weight distribution through sliding elements (Zenner and Krüger, 2017) and thus trick the somatosensory system into believing an object is gaining or losing weight, when it is really only getting more difficult or easier to handle. The actual *mass* of an object is conceptually a variable too, and would mostly appear as *weight* unless the haptification happens in a micro-gravitational environment. It might be a rare occurrence in dynamic systems, however one could for example imagine a larger moving object that is not worn or carried and filled with additional material as a data attribute increases.

Griffin's variables of resistance and friction on the other hand are very focused on haptic maps, in that friction is specifically for resistance encountered moving across or through a virtual surface, while the variable of resistance is for how resistant a material is to deformation. In virtual environments with arbitrarily shaped fully virtual objects and potentially complex volumetric structures, this distinction is difficult to uphold, as unless some tangible object is present, both are the same simulated resistance. More appropriate for virtual environments perhaps are considerations by Harding et al. (2002), which talk about conceptualizing haptic volumes as force fields. They explicitly name the variables of *attraction* and *repulsion*. Anything that resists just by "being in the way" can then be grouped under a *resistance* variable.

With resistance, attraction and repulsion defined through a force field analogy, what about friction? In physics simulations, friction tends to have one of two meanings, depending on context—surface friction in a material (making it a tactile variable, or as we will see later, a variable of an object surface), or the air resistance within a volume (i.e. the resistance of a volume against movement within it, not the resistance of an object at its borders). Surface friction is arguably closer to the everyday understanding of friction, which is why it will get the name of *friction* in the variable space, while air resistance causes *drag*, which is quite compatible with the previously established force field metaphor.

Because we have tangible objects in our space of possibilities, there is also a need for a variable describing their *material* properties. This is any property that goes beyond active movements in the object and is not already covered by the general texture variable. Recent advances in material science make this class of variables more broadly attainable for use in interactive systems. There are many properties of materials defined in the realm of physics, specifically relevant here are: elasticity, plasticity, viscosity, stiffness, hardness, toughness, and malleability. There are also properties that concern the surface of a body that are not a result of active tactile systems, like roughness, slickness, surface tension and wetness (Tak and Toet, 2013). Thus, we should look at the tactile variables from earlier as "active tactile" variables, while here we refer to static surface features and textures. It should be noted that even with current material science technologies, few of these basic properties are dynamically changeable through digital inputs. However, they can often still be implied through other haptic variables or sensory illusions induced by cross-sensory effects (pseudo-haptics). It is important to remember that we do not always have to manipulate physically existing objects, but we can also create purely virtual haptic objects that we try to model as close to a real object as possible.

Finally, a distinction needs to be made between the active creation of temperature as imagined in Griffin's variables and the potentially changeable material property of *heat conduction*. Two materials of the same temperature degree can feel quite different to the human touch, as our thermoreceptors really measure thermal conduction instead of absolute temperature. To reduce some of the complexity, we treat all the material properties of objects as one variable, and the tactile properties of their surface as another, similar to the texture or composition variables from before.

The final haptic variables thus are:

Sp1: Location The location of an objected, discovered once it is touched. In order to encode data attributes in this, small changes have to be felt over time or in relation to other reference objects.

Sp2: Size The physical size of an object, as discovered through touch. Can operate in the same capacity as visual size, for example for a kind of haptic cartogram.

Sp3: Orientation Because location and shape of an object can be felt, feeling orientation and changes in it is equally possible.

Sp4: Shape When manipulating an object that changes shape, haptics is perhaps uniquely qualified to make a user able to feel complex topological changes.

H1: Vibration intensity How forcefully a surface or object vibrates against the skin.

H2: Vibration frequency Similar to audio signals, the frequency of the vibration can be felt on the skin. This can range from lower-frequency oscillations that create a rhythmic thumping, to high-frequency vibrations that make an object unpleasant to the touch.

H3: Vibration envelope How abruptly or softly each vibration pulse hits.

H4: Air flow The sensation of air blowing against or past the surface of the skin.

H5: Temperature Heat radiation hitting the skin and warming it. This radiation has no way to induce an "opposite" signal, the skin has to cool down on its own.

H6: Pain Pain sensations induced within the surface of the skin.

H7: Kinesthetic alignment The difference between actual proprioception and what the virtual world is showing. Requires a simulated virtual body or at least some form of body-analogue.

H8: Attraction Describes force that is applied towards an object or a certain part of an object. This could for example lead a user towards a specific value in a field of values.

H9: Repulsion The opposite of attraction.

H10: Resistance How difficult it is to move an object that is fixed in space out of its fixed position. This is often a feature of buttons and triggers, which offer a certain amount of resistance before they give in and can be moved out of position.

H11: Drag Any resistance that is applied specifically when the user is moving—as opposed to attraction and repulsion, no force is applied when no movement occurs.

H12: Mass The actual mass of an object—this can only be modulated by actually adding material.

H13: Mass distribution Where in an object its mass is concentrated, while the total mass remains the same.

H14: Material Internal properties of the material of an object, including: elasticity, plasticity, viscosity, stiffness, hardness, toughness, malleability, heat conduction.

H15: Surface External properties of the surface of an object, including: roughness, slickness, friction, surface tension and wetness.

3.1.4.4 Olfaction

Even though it is usually not seen as one of our most important senses, olfaction is without a doubt the most complex in the "signals" it resolves. It senses the complexities of odor molecules, which allows it to distinguish thousands of distinct chemical stimuli (Strugnell and Jones, 1999). It is used to detect chemicals carried into the oral and nasal cavity, either through the air or through direct contact, in order to monitor what we consume or inhale for palatability, danger, and familiarity. Perceptions of smell are deeply linked to memories and emotions (Marriott et al., 2018), and are thus highly subjective. However, olfaction objectively has an outsized positive impact on memory-related tasks and is arguably a heavily under-explored modality in computing (Garcia-Ruiz et al., 2021).

There are large gaps in our knowledge about the dimensions and attributes of olfactory perception. The sense at its core is one of two senses that detects chemicals, gustation being the other. But while we distinguish between only a handful of dimensions of taste, olfaction is highly multidimensional. Study subjects often have problems even verbalizing what kind of smells they are experiencing—usually they will be related to one or multiple guesses for a source substance, instead of direct descriptions of perceptual attributes (Marriott et al., 2018). We can describe the colors of a flower or the softness of its petals, but we can not describe its distinct smell without referring to the flower itself. There also seem to exist complex interrelations between chemicals, which makes this sense even more complex than just a list of several thousands of chemicals (Strugnell and Jones, 1999).

A discipline that is obviously interested in possible categorizations is the field of perfumery. Their categorization needs are somewhat different than the ones needed for data olfaction, however we will still consider them in the course of this section.

Hardware

Olfactory displays roughly fall into two categories: displays which use some way to physically move odorants towards the user's nose through the medium of air (both at long and very short distances) and displays that are inserted into nose or mouth for direct contact. For obvious reasons, the former tends to be more popular, both in stationary as well as worn forms.

The main limitation of all olfactory displays is that there are no basic components to break a complex smell down into—it is just complex on a chemical level. In practice this means that olfactory displays only come with a certain selection of odorant cartridges, whose odors can then be freely combined. The selection of odorants is done according to the application. While an application that just needs a good, a neutral, and a bad smell to conceptually highlight data facts could use many different odorants and even switch them out between uses, an application that tries to portray the "smellscape" of a modern building would need a highly specialized mixture of fitting odorants.

The way these odorant cartridges work is akin to perfumes. Their contents usually consist of a mix of alcohol, distilled water and the odor substance itself (essential oils or synthetic substances), which are then dispersed into the air as fumes and blown into the desired direction. Sometimes during that dispersal the odorant can be further diluted at varying rates, to modulate the intensity of the smell as it arrives at the user.

Another way to move odors is the introduction of scented or naturally smelling objects into the environment. This however is not a very targeted method and makes it hard to modulate intensities or stop the introduction of a smell. A way to do this more dynamically is to burn the scent objects, like in the case of candles. There are also systems where scratching a surface releases the scents stored in a small film of material, which could conceivable be combined with tangible objects and haptic interactions. A very different category of smell display are contact smell displays that are currently in their early stages. These displays do not utilize odorants, but try to directly stimulate the olfactory receptors in the nose and mouth through electrodes (Cheok and Karunanayaka, 2018).

Variables

For the variables, we have to grapple with two problems: First, the highly multidimensional space of different odor perceptions. It is as impractical to encode every possible smell into a variable as it is impossible—our understanding of olfactory biology is simply not complete at the moment. Secondly, the multiple forms of vulnerability to sensory error when encountering a complex array of odors. There are several thresholds for an odor signal, which can be different in each person: The point at which it becomes perceptible at all, then the point at which it becomes clearly distinguishable, and finally a point at which the odor become too strong to comfortably stay in the environment. Then there are the odorant molecules themselves, whose vibration, weight, shape, chain length, and more determine how fleeting it is (both in the air and on the skin/nerves) and how well it covers other scents that are already in the environment. Even knowing all of the molecular properties of a substance, it is still very difficult for us to predict its smell profile (Kaeppler and Mueller, 2013).

Modern parfumery tries to circumvent the first problem by employing the so-called "fragrance wheel" (see Figure 3.2), which derives from historical categories of perfumes that have been expanded with new synthetic capabilities over time. It tries to be a tool akin to color wheels, on which concepts like complementary smells can be employed. And while it has some explanatory power for the specific context of perfumes and fragrances, it does by definition not include any sort of smell that would not be part of what is historically and culturally appreciated as body or room odors. It is, like many other smell taxonomies, simply a way to verbalize and talk about odors in a comparative manner. No objective taxonomies exists. Even after over a hundred years of odor science there is little more than a collection of incomplete approaches and falsified categorizations (Kaeppler and Mueller, 2013).

FIGURE 3.2: Michael Edward's fragrance wheel, from Edwards (2019).

The second problem is also tackled by working from experience and categorizing fragrances into different kinds of notes: base notes, middle notes and head notes. It is very difficult to be precise about this, because this musically-inspired model simplifies several effects, including how long odors linger (persistence), one odor enhancing others (multiplicity), and how well one odor suppresses another (masking) (Garcia-Ruiz et al., 2021). Additionally, humans are also capable of quickly growing accustomed to ambient smells, moving them from conscious perception to a subconscious level. All this is further complicated by subjective effects, like the deep cultural and personal entanglements of odors caused by the tight interrelation of olfaction with our memory apparatus.

With all this in mind, olfaction is by far the least robust sense to employ for any sort of controlled data sensification exercise. Despite these difficulties, we need a good approximation of variables, in order to even begin to tackle these issues. Tak and Toet (2013) off-handedly note the two variables *intensity* and *valence*, but do not explain what is meant by valence exactly.

We combine this with an approach also used by perfumers themselves: adapting vocabulary from music. In this current context, the equivalent would be the sonification variables. Volume is simply the *intensity* of a scent representing some data fact. Pitch is the smell of the used odorant, or of the main odorant in a larger composition. In line with chemo-sensory research done mostly in the taste domain (Kelling and Halpern, 1988), this will be called the smell *quality*. Timbre describes the *layering* of many different qualities into one smell. The envelope is also encoded in this same variable, as the base notes will for example be perceptible longer than the head notes. This factor however is not separable like it is in audio signals.

All other temporal factors are best described by the time variables—removing or masking certain scents on the fly is however still a challenging problem in olfactory displays (Sugimoto et al., 2010). Clarity variables like noise and quantization might technically be possible, however the robustness of smell is not sufficient and our smell perception likely not precise enough for us to effectively perceive minute details and defects like we do in vision and hearing.

Another set of variables other senses bring with them are the spatial variables location, size, orientation and shape. However, smells spatial resolution is arguably not high enough to expect any spatiality but simple location. Location at least can be determined by "sniffing around", though the usefulness of this variable in olfaction is very context-dependent.

Sp1: Location Olfactory objects can be detected over time, by moving and inhaling in a space.

O1: Intensity The strength of an odor.

O2: Quality The smell sensation as created by one pure odorant or by the dominant odorant in a more complex odor.

O3: Layering The entire composition of an odor, with all its layers and the natural change in odor perception the user experiences over a short amount of time.

3.1.4.5 Gustation

The gustatory sense stems from chemical receptors (taste receptors) that are located on taste buds. The taste buds themselves are distributed all over the oral cavity, but mostly focused on the tongue, where they are spread over little bumps called papillae. Functionally, taste is the sense that detects how nutritionally valuable a consumed substance is (Lindemann, 2001). Every one of the main five tastes (sweet, salty, sour, bitter, and umami) is associated with a certain class of nutrients. Sweet with carbohydrates, umami with protein, salt with sodium and other ions, sour with acids, and bitterness with potential toxins and plant alkaloids (Breslin and Spector, 2008). For our purposes however, perhaps the most important part in defining the sense of taste, is to define what it is not: It is not the sense that covers all intricacies of food, drink, and other substances that enter the oral cavity. Instead, it is just a small part of the whole, mainly covering the five tastes. The olfactory sense is responsible for most of the chemical complexity of flavor and the haptic sense for the sensations of texture, temperature, pungency, wetness, and more. This sensory integration is usually referred to as the *flavor* of a substance instead of taste, and represents perhaps the most common example of complex multisensory integration inherent in the body.

Hardware

Taste is possibly the most invasive primary sensory modality to display to. Any display has to at least be inserted into the mouth or even chewed. Currently, the most common display format are the gustatory equivalent of tangible objects: edible data objects (Mueller et al., 2021). This has several problems, in that these objects are difficult to create on the fly as a reaction to interaction, in the way we would for the other senses, but they are also more fleeting than normal tangible objects in that they usually get destroyed every time they display their taste. A way around this could be objects that display taste when licked, however user onboarding for such systems might be particularly difficult.

Perhaps more palatable are electrode-based direct simulation technologies, similar to the ones currently in development for olfaction (Cheok and Karunanayaka, 2018).

Variables

The first concern in terms of variables are the five main tastes. Conceptually they are very similar to color perception: There are multiple different receptors that are co-located and have a focus point (the fovea and the tongue respectively). Color is usually separated into three values, either red, blue, and green, or saturation, hue, and value. The obvious question then is whether it is possible to construct something akin to a color wheel for the five tastes. This is unlikely, as taste and vision diverge in several key places: Apart from just the dimensionality (three colors as opposed to five tastes), the degree to which there is metamerism (different combinations of physical qualities producing the same perceptual result) within the currently known aspects of the sense of taste is a topic of debate (Breslin and Spector, 2008). Without metamerism, the five-dimensional taste space can not be geometrically collapsed into a more useful tool in the same way the 3D color space is turned into a wheel. If there was such a way to construct a simplified geometrical representation of taste, it would likely still be far more complex than the simple spectrum we are used to in color perception.

For this reason, there are different encoding systems in different disciplines that deal with taste. The one that matches most closely our way of describing color is the viewpoint of chemosensory research that we already employed for olfaction, in which taste is measured in taste quality and taste intensity (Kelling and Halpern, 1988). Taste *intensity*, like value and saturation, is conceptualized as "prothetic" perceptual dimension, in which values can be arranged from less-to-more (Spence, 2019). Taste *quality* is, like color hue, a metathetic continuum, in which this is not possible. In the sensification language we have adapted from Bertin (1967) and MacEachren (1995) this would mean that taste intensity is useful for ordinal values, while taste quality is useful for nominal values.

The profiles we can build purely in the domain of taste are a lot less complex than what is possible with other senses—most of the complexities we are used to in our food only exist in reference to flavor. It thus does not make much sense to include a timbre or texture-style variable here. The taste in our mouth at any point is thus actually well encoded by just a handful of quality dimensions and an intensity.

A temporal sensation that is specific to gustation is the concept of *aftertaste*, in which a taste keeps developing in complex ways even after the tasted substance has left the mouth. Because aftertaste is something that can be modulated by introducing different substances into a recipe, it could conceivable be used as a sensory variable.

Taste is also the only primary sense that does not permit spatial variables—all explorations of an objects form or location in the mouth would instead be caused by the haptics of the mouth, not its taste receptors.

Beyond the commonly known qualities of taste, there still remains scientific discussion about further sub-modalities. These cover possible additional taste receptors for calcium and lipids (which would be part of the quality variable), as well as the sensations of *pungency* (as in spicy food) and *coldness* (as in minty flavors). Spiciness and coldness however are more likely to be "misfires" or even intended secondary uses of somatosensory receptors for temperature perception, as they can happen even in parts of the body that do not have taste receptors (Lou, 2012). These aspects are thus better represented as either part of taste quality, temperature (H5) or surface (H15), depending on the specific context.

G1: Intensity The magnitude of a taste impulse.

G2: Quality The relative mixture of taste sensations (sweet, salty, sour, bitter, and umami) included in a taste impulse, possibly extended by somatosensory sensations like pungency or coldness as they appear in the mouth.

G3: Aftertaste How the taste impulse changes once the edible substance has left the mouth.

3.1.4.6 Secondary Senses

The secondary senses are also called the "interoceptive senses", as they are the senses that offer us information about the current state of our body, our internal organs or ingested substances within our body. The only primary interoceptive sense is proprioception, which was already discussed in the haptics section. In haptics, we mostly considered it as a variable of misalignment—for example by showing data about the environment by making the user's virtual body sluggish. If we take the interoceptive conception, there might be ways to manipulate the user's proprioception in itself. This would be represented by a variable called *kinesthetic manipulation*.

There is some debate about which other secondary senses exist and how they function. Commonly named examples are: Sense of balance and spatial orientation (vestibular system), nociception (for non-haptic pain sensations), hunger, thirst, respiration, suffocation, nausea, salt and sugar levels, blood flow, blood pressure, gastrointestinal stretch, swallowing or vomiting, acid reflux, fullness of bladder and rectum, and heart activity. There are also often hints for non-conscious, low-quality sensing of sensory inputs that are at the edges of what our primary senses can detect, like ultraviolet and infrared perception or very low frequency sound wave perception.

The real question is: Can or should any of these actually be used to encode data? Two interoceptive senses that are already important (and frequently being manipulated) in immersive applications are the sense of balance and spatial orientation. There has been much work on tricking them (Steinicke et al., 2008) as well as some work on more direct manipulation, especially in how they pertain to the issue of motion sickness (Miller and Muth, 2004).

In the future, further interoceptive senses could conceivably be fooled or directly manipulated to highlight certain usually non-conscious body-internal processes, possibly for medical applications. This is however unlikely to become very relevant for geospatial applications outside of highly experimental interfaces.

One possible exception could be non-conscious fringe sensations like extremely low or high frequency sounds. Once we have better evidence on what changes in behaviour such non-conscious signals might cause, they could be used to communicate very subtle, complex shifts in a virtual world or in the underlying data, perhaps introducing forms of unease or relaxation as required during interactions.

For this thesis, the only other secondary senses variable we will include is the variable of *balance*. The sense of balance is related to the proprioceptive sense, but integrates multiple different organs like eyes and ears. It is easily fooled by external methods like moving seats or platforms, or even more invasive methods like TENS (transcutaneous electrical nerve stimulation) (Lee et al., 2022). We thus have the following final selection:

P1: Kinesthetic Manipulation **P2: Balance**

3.1.4.7 The (Spatiotemporal) Multisensory Variable Space

After all this, we adapt MacEachren (1995)'s visual variable syntactics for the multisensory space of variables proposed in this section. MacEachren's original tool was an extension of multiple attempts to categorize variables, first introduced by Bertin. With five different categories of effectiveness, it is the most detailed form of syntactics in the literature.

A variable is useful for values with *nominal* scales if we can distinguish between multiple different classes, without these classes having to be ranked in any way. A variable works well for values with *ordinal* scales if we can use it to rank two or more signals of different strengths. A variable works well for values with *numerical* scales if we can accurately read or at least guess reasonably well at what number on a scale it represents, usually given some form of legend.

Different from how effectively these three kinds of value scales can be represented, is a variable's perceptual efficiency in relating many values at once. After Bertin, a variable is useful for *associative* analysis if it enables us to quickly group similar values together, without giving certain values precedence. *Selectivity* describes how well certain values can be picked out of a large pool of values, even with symbols that are encoded with multiple variables. MacEachren calls these factors visual levels (associativity) and visual isolation (selectivity). For our classification, we will stick with Bertin's original names, as they are more sense-agnostic.

At this point we need to acknowledge one more aspect that MacEachren et al. already mentioned when writing about these concepts: certain variables are analogous to each other over multiple senses. "location with location, loudness with size, pitch with value and timbre with shape" (MacEachren, 1995). The context for this was of course very different, so the analogues will not be the same, but nonetheless they are there in the immersive case. One way in which such analogies can be found are to compare how they map a value over an object, i.e. whether they operate similarly in a spatial sense. Another way is to compare whether they are similar in how effective they are in representing certain scales or how they act on the viewers perception.

What remains now is to apply these syntactics and the concept of variable analogies to our variable space. The results of this adaption process are shown in Table 3.1. It brings all sensory variables into one taxonomy, lists which variables are analogous over what senses, which types of values can be effectively displayed, and what the perceptual attributes of the variable are. It also tries to mark existing judgements from the literature where possible—at least as far as they carry from the non-immersive to the immersive case.

Just like in the literature, most of this is based on hypothesis or informed opinion instead of empirical studies. It should thus be taken as a guideline, parts of which might change as we learn more in the future.

Variable Type	Variable	Sensification Feasibility						Scale Representability			Perceptual Attentiveness	
		V	S	H	O	G	I	Numerical	Ordinal	Nominal	Associativity	Selectivity
Temporal	(T1) Duration	x	x	x	x	x		(M/Ko)	(M+Ko)	(M)		
	(T2) Rate of change	x	x	x	x	x		(M/Ko)	(M+Ko)	(M)		
	(T3) Order	x	x	x	x	x		(M)	(M+Ko)	(M)	(Ko)	(Ko)
	(T4) Display date	x	x	x	x	x		(M)	(M)	(M)		
	(T5) Frequency	x	x	x	x	x		(M)	(M/Ko)	(M)	(Ko)	(Ko)
	(T6) Synchronization	x	x	x				(M)	(M)	(M)		
	(T7) Quantization	x	x									
	(T8) Composition	x	x	x	x	x						
	(T9) Agent behaviour	x	x	x	x	x						
Spatial	(Sp1) Location	x	x	x	x			(M)	(M)	(M)	(M)	(M)
	(Sp2) Size	x		x				(M)	(M)	(M)	(M)	(M)
	(Sp3) Orientation	x		x				(M)	(M)	(M)	(M)	(M)
	(Sp4) Shape	x		x				(M)	(M)	(M)	(M)	(M)
	(Sp5) Environment	x	x									
Linguistic	(L1) Speech		x									
	(L2) Writing	x		x								
Appearance	Color value	V1						(M)	(M)	(M)	(M)	(M)
	Color saturation	V2						(M)	(M)	(M)	(M)	(M)
	Color hue	V3						(M)	(M)	(M)	(M)	(M)
	Texture	V4						(M)	(M)	(M)	(M)	(M)

TABLE 3.1: A list of sensory variables by type. "x" marks what sensification modalities each variable is available to. For scale types and perceptual qualities the grayscale value denotes how effective a variable is for that use case. Dark grey is effective, grey is possibly/marginally effective or effective if interactive, white is not effective. Where applicable, it is noted what authors have made this or a similar judgement before. (M—(MacEachren, 1995), K—(Krygier, 1994), Ko—(Köbben and Yaman, 1995), G—(Griffin, 2001). If multiple authors assign a value, a "+" indicates agreement, and a "/" disagreement.) Colors mark current technical feasibility of the required display systems: Available commercially, Feasible to build, Challenging or impossible to build

		Sensification Feasibility						Scale Representability			Perceptual Attentiveness	
Variable Type	Variable	V	S	H	O	G	I	Numerical	Ordinal	Nominal	Associativity	Selectivity
Display	Blur	V5						(M)	(M)	(M)	(M)	(M)
	Transparency	V6						(M)	(M)	(M)	(M)	(M)
	Level of detail	V7						(M)	(M)	(M)	(M)	(M)
	Culling	V8										
Spectral	Volume/Intensity		S1	H1					(K+G)	(K+G)		
	Pitch/Frequency		S2	H2					(K)	(K)		
	Timbre		S3						(K)	(K)		
	Envelope		S4	H3					(K)	(K)		
	Noise		S5									
	Compression		S6									
Tactile	Air flow			H4					(G)	(G)		
	Temperature			H5					(G)	(G)		
	Pain			H6								
Kinesthetic	Alignment/Manipulation			H7			P1		(G)	(G)		
	Attraction			H8								
	Repulsion			H9								
	Resistance			H10					(G)	(G)		
	Drag			H11					(G)	(G)		
	Balance						P2					

TABLE 3.1: A list of sensory variables by type. "x" marks what sensification modalities each variable is available to. For scale types and perceptual qualities the grayscale value denotes how effective a variable is for that use case. Dark grey is effective, grey is possibly/marginally effective or effective if interactive, white is not effective. Where applicable, it is noted what authors have made this or a similar judgement before. (M—(MacEachren, 1995), K—(Krygier, 1994), Ko—(Köbben and Yaman, 1995), G—(Griffin, 2001). If multiple authors assign a value, a "+" indicates agreement, and a "/" disagreement.) Colors mark current technical feasibility of the required display systems: Available commercially, Feasible to build, Challenging or impossible to build

Variable Type	Variable	Sensification Feasibility V	S	H	O	G	I	Scale Representability Numerical	Ordinal	Nominal	Perceptual Attentiveness Associativity	Selectivity
Physics / Simulation	Mass			H12								
	Mass Distribution			H13								
	Material			H14								
	Surface			H15					(G)	(G)		
Chemical	Quality				O1	G1						
	Intensity				O2	G2						
	Layering				O3							
	Aftertaste					G3						

TABLE 3.1: A list of sensory variables by type. "x" marks what sensification modalities each variable is available to. For scale types and perceptual qualities the grayscale value denotes how effective a variable is for that use case. Dark grey is effective, grey is possibly/marginally effective or effective if interactive, white is not effective. Where applicable, it is noted what authors have made this or a similar judgement before. (M—(MacEachren, 1995), K—(Krygier, 1994), Ko—(Köbben and Yaman, 1995), G—(Griffin, 2001). If multiple authors assign a value, a "+" indicates agreement, and a "/" disagreement.) Colors mark current technical feasibility of the required display systems: Available commercially, Feasible to build, Challenging or impossible to build

3.1.4.8 Cross-Modal Effects

An important aspect not included in Table 3.1 is the interrelation of variables. Cross-modal effects between the different senses and the complexities of such overlaps in different situations and environments are simultaneously a subject of a large amount of psychological research, as well as drastically understudied. The reason for this is the sheer complexity of the problem: The perception of every variable could interfere with every other variable, depending on the location that they are displayed from, exactly how they are displayed and the environmental conditions in which they are being displayed. As such, a complete picture of these effects would require an enormous number of user studies, making it impossible to generalize them in their totality. However, some important results and general principles should still be kept in mind in mind.

For the conventional visual variables in the cartographic context there is for example Roth (2017), who examines "variable conjunctions and bivariate mapping" strategies. He identifies that such mappings can either try to strengthen the display of one attribute through redundancy, or map multiple attributes into one symbol/location through multiple variables. Multiple examples of such mappings are shown that conform to common cartographic visualization strategies. Like the visual variables themselves, this is based mostly on experience and partly informed by psychological research. A similar but empirically validated approach was taken in Ogi and Hirose (1997) for the multi-sensory case over the variables color, loudness, sound frequency, and air flow pressure.

Even though this strategy of constructing examples and discussing them in depth is too complex for a variable space as large as the one established in this chapter, Roth highlights an important aspect of this topic: variable redundancy and multi-variable mappings are two very different sides of the same coin. At first glance, it might seem that redundancy will always be an advantage, while the real trouble lies within multi-variable mappings that start to interfere with each other. However, as Kapralos et al. (2017) write, redundancy can also unnecessarily increase our cognitive load and exceed sensory channel limits in such a way that it ends up lowering effectiveness.

Still, multisensory integration is a critical part of our cognition in real-world environments (Wesson and Wilson, 2010), and in many cases multi-modal interfaces end up increasing our sense of immersion and presence in virtual environments (Kapralos et al., 2017). Roberts and Walker (2010) marks the use of cross-modal effects as one of the five main aspects of multisensory information visualization. The most obvious advantage for embodied geosensifications specifically would be the possibility of encoding different attributes of one geospatial feature over multiple senses, aiming every attribute precisely at the sensory variable that is most fit to perceive it.

3.1.5 Conclusion

Throughout this section, we have considered an interdisciplinary body of literature concerning visual variables and how they have been extended from the visual domain into the other senses. We have established the categorizations of visualization, sonification, haptification, olfaction and gustation as they will appear in the rest of the thesis. Most importantly, we have derived a multisensory space of variables for representation of data values in immersive and situated geographic scenes, complete with rankings for their capability to represent scales, their perceptual qualities and technical feasibility.

However, as established before, the sensory is only one aspect of sensifications. Just like visualizations, sensifications depend on the fidelity of the data that they are supposed to display and the interactions we are able to use to investigate that data. Regardless of the progress made in the space of sensory variables in this chapter, these mappings alone do not allow us to specify the full spectrum of what a sensification is. In order to be able to do just that in Chapter 4, we now need to investigate both data representation as well as interactivity in geographic scenes.

3.2 Data—Representing Geospatial Objects

"Maps are powerful tools, and have been for centuries, because they allow us to see a world that is too large and too complex to be seen directly. The representational nature of maps, however, is often ignored—what we see when looking at a map is not the world, but an abstract representation that we find convenient to use in place of the world."

Alan M. MacEachren, How Maps Work (MacEachren, 1995)

One of the defining factors established in Chapter 2 was the difference between general visualizations and the specific use case of geospatial data. The reason this divide is so notable starts at a very low level: the raw data. For geospatial visualizations, most of this raw data is going to be GIS data, meaning that it will be referenced to some global coordinate system and consist of either raster or vector data structures encoded in either a textual or binary format. Because this data has such high practical applicability, there is a basic coordinate referencing process all proper GIS data follows, as defined in the 191xx family of ISO (International Organization for Standardization) norms. In order for the practical considerations as well as the concepts developed in this thesis to stay relevant to everyday applications, we will have to stay close to these basic realities. At the end of this chapter, what this should yield is a taxonomy of the kinds of geospatial features that we can *communicate* facts *about* and that we can *interact with* in embodied immersive contexts.

GIS data today comes in more formats than one person can keep track of. The same basic feature can often be represented by completely different data structures without any or with very little loss in accuracy. The height of a landscape for example could be saved as an image in a raster format, as a regular array of 3D points or as a triangulated irregular network (TIN). Data will be saved with either 2D, 2.5 or 3D coordinates. Some features will not be saved as geometry at all, but encoded into attribute values, like a height value in a tree data set denoting the 3D height in relation to the 2D point feature. (Bill, 2023)

If these classifications were sufficient for embodied geosensification, we could stop here. The intricacies of raster data and vector data, of points, lines, and polygons and the different data formats have been discussed many times, and by now are a simple reality of all spatial data (pre-)processing, in many cases self-explanatory enough to not warrant explicit mention in shorter publications. However, to call back to Lü et al. (2018), immersive environments are different. Suddenly, organising GIS data by layers or feature objects is not sufficient anymore. What does it mean for a tensor field to be geospatially referenced or for the current cloud cover over a region to be available as the sort of signed distance field that a 3D engine could render efficiently? What do we do if a bridge in our city model needs to sway in the wind and thus requires a rigged skeleton? Spatial data file formats are not prepared for this today, often requiring complex, non-standardized, and lossy conversion steps between georeferenced data formats like Shapefile, scene formats like GLTF (Graphics Language Transmission Format), and the many specialized encodings used for fields, materials, and bodies in computer graphics and visual effects. And while all this is true for encodings and formats that have been around for decades, radically new methods of data representation are already on the horizon. One example are the neural network-driven neural radiance fields (NeRF) first proposed in Mildenhall et al. (2021), which have taken volumetric rendering by storm. New neurally encoded formats could go as far as to shatter our conceptions of what data encodings (or even data itself) has to look like.

There is no way to foresee what (spatial) data for immersive environments will look like in ten or twenty years time. The only reaction that remains is to try to stay as encoding-agnostic as possible. Usually when unsure about the foundations of a certain topic, it pays to go back in time and see what conceptualizations there were before certain technical decisions had become accepted and commonplace. In this case, this means to trace back the genealogy of the representation of geospatial objects. Geospatial objects in this context shall be any past, present, simulated, imagined or planned features and phenomena of the real world, as well as any facts about the world, that can be assigned some sort of global position and usefully stored on a computer. The complexities of practical GIS data will only be discussed where required.

If we consider this definition carefully, we can already see that the common idea of "position" is always an abstraction. We happily assign objects point coordinates, even though they are volumetric objects in the real world, making their point-position arbitrary at a certain scale. Facts about the world can be highly distributed and chaotic and our knowledge of them fuzzy, and yet we try to store them in a perfectly regular raster.

Most commonly associated with the foundations of spatial data is perhaps the geographer Michael Goodchild, who with his colleagues lays out a comprehensive concept for a general theory of geospatial data representation in Goodchild et al. (2007). Although this is far from being the most active field of research, new work is being done to this day. Examples include the work on events and time series in Ferreira et al. (2014) and on the realities of observation and sensing in Hu et al. (2020).

At the base of much of the debate over these representational models lies a distinction between two different representational strategies, often seen as mutually exclusive. The most common incarnation of this distinction is the difference between raster and vector data, or also "maps-as-images" and "maps-as-geometric-structures" (Peuquet, 1988), which every user of GIS will very quickly have to grow accustomed to. Goodchild and others put this distinction on a more abstract, mathematical foundation: the duality between continuous and discrete representations, or *fields* and *objects*. Couclelis (1992) writes: "The ongoing debate in GIS regarding the relative merits of vector versus raster representations of spatial information is usually couched in technical terms. Yet the technical question of the most appropriate data structure begs the philosophical question of the most appropriate conceptualization of geographic space", and goes on to utilize the same field-object distinction as Goodchild, however for perceptual reasons. She argues that we perceive and understand some of the phenomena around us as discrete objects (a tree, a building, etc.) and some of them as continuous fields (gravity, terrain, environmental noise, etc.).

Goodchild et al. (2007) deviates on this last point and argues that most phenomena, if seen and measured at a single point in time, could cogently be described as both fields or objects. It depends on the individual application which might work best. If we want to count individual cloud formations, we want to see clouds as individual objects, if we want to measure risk of rain we might look at cloud formations as continuous and shifting. Both of these authors also specifically consider this distinction as a way to approach data uncertainty and error (Goodchild, 1989; Couclelis, 1996).

In the context of this thesis there is also another reason to not talk about the representation duality in terms of vector and raster data—only at a certain resolution is raster data continuous and vector data discrete. Once we move closer to the data, say to a 1:1 scale with an egocentric viewpoint, we quickly realize that the relationship seems to flip: a raster suddenly appears as a discrete way of storing the data—any point that holds relevance is exactly localized by a simple pair of whole numbers, locked in a grid that is too coarse to appear real. In vector data on the other hand, individual points of a geometry can have arbitrary positions in the real number space, thus the number of points that lie inside a vector polygon are infinite and continuous. Vector data, especially when curves and splines get involved, is often just as much about the geometries that connect individual points as it is about the points themselves. Once we find ourselves standing on very obviously blocky, procedurally generated terrain in a virtual environment, we wish that the elevation model underlying it had been mapped in curves between distinct points instead of a "continuous" raster.

Considering all this, it only makes sense to steer towards the *field vs. object* view of geographic representation. The question then is: How do we, in the face of the often interchangeable nature of fields and objects, build a working, usable theory on this basic distinction so that it does not become arbitrary. Here, Goodchild et al. (2007) builds on previous work for example from Kjenstad (2006) and introduces the *geo-atom*. They define a geo-atom as "an association between a point location in space–time and a property", represented by "a tuple $(x, Z, z(x))$ where x defines a point in space–time, Z identifies a property, and $z(x)$ defines the particular value of the property at that point." The dimensionality of the point x implies the kind of space we are working in. A simple non-temporal 2D data set just needs a point with two coordinates, a spatiotemporal 4D application needs four coordinates, three in space and one in time, while a space fully constrained to one string of lines would need just one dimension. The notion of field and object then comes into play through the function z. For objects, the function is only defined where it is within or on the border of one or multiple geometries, for fields it is a topologically continuous function defined throughout the whole coordinate space. Every set of GIS data is in principle reducible to a set of these geo-atoms, although that set may be of infinite cardinality. Goodchild et al. (2007) then defines every statement about higher-level geographic objects as being the result of some sort of aggregation of geo-atoms based on a rule—a statement about a line segment for example could be made by first aggregating all geo-atoms that fall on a line between two points. If we imagine a collection of such rules we might get a result that looks close to an attribute table of a point dataset in GIS software: an (infinitely long) list of coordinates with multiple properties Z attached to them.

From this geo-atom and the possibility of aggregating them, Goodchild et al. (2007) then establishes geo-fields and geo-objects. A *geo-field* is an aggregation of one (for scalar fields) or multiple (for vector fields) properties Z over a domain D (a subset of the vector space in which all the geo-atoms are defined). A *geo-object* is an aggregation of points whose geo-atoms meet certain requirements in one or more property values. An individual building for example could have a building ID and every geo-atom that has its property "building id" set to the correct number is part of this geo-atom. Goodchild calls these measures that define geo-objects and geo-fields *spatially intensive*, because every constituent geo-atom needs to carry them. Measures that can only be defined over the whole of an object or field, such as its area or volume, are called *spatially extensive*.

Interestingly, the rift between these two definitions is easy to cross. Goodchild et al. (2007) posits that for any geo-object we could define a membership-function $m(x)$ that takes the value functions $z(x)$ of one or multiple properties Z and calculates from them a measure of membership to a geo-object for each point in space. This makes every geo-object reducible to a geo-field and even allows fuzzy objects through partial membership. By a similar mechanism we could also define objects with internal complexities, like a storm formation being bounded by its cloud cover, but having multiple internal stages defined by pressure and precipitation, like the famous eye of the storm. Yuan (1999) calls this “field objects”.

In reality we can only ever sample a limited number of points from this space, because we do not possess infinite storage, measurement accuracy, and computing capabilities. We thus need to "discretize" these structures. That this is possible while still retaining useful data is due to the high spatial *autocorrelation* within geographic phenomena—a property like elevation is not expected to change to any meaningful degree over any given infinitesimal distance. Goodchild et al. (2007) also goes further into other practical matters, like zonal operations as commonly found in GIS systems, where we take properties that should technically be spatially extensive and make them intensive, which then necessitates dealing with how to split and merge these fields and objects.

What is most important for immersive visualization however, is that we expand the conception of what a geospatial object could be. It is no accident that the example of clouds or storm formations was mentioned multiple times: a GIS data format fit for immersive environments should be able to store a cloud—a volumetric object with fuzzy borders that encodes multiple continuous value fields within itself, and changes in complex ways over time, possibly even merging with or splitting into multiple other clouds. For the geometry and value aspect of this problem, the concepts discussed so far offer us the necessary theoretical basis.

Goodchild’s work also discusses time as a factor beyond just its conception as another dimension of change. Of specific interest here are geo-objects, which can change over time in multiple unique ways: they can be moving or stationary in position, uniform or evolving in internal structure (i.e. changes in their internal value fields), and elastic or rigid in geometry (as in changing their shape). This modulation is directly linked to how well we can measure certain phenomena—tracking changing shapes or internal structures remains a rarity in most common GIS use cases. This level of fidelity is exactly what is often missing in immersive applications.

Ferreira et al. (2014) builds on top of certain aspects of these theoretical frameworks, by explicitly focusing on the temporality of data. They start at the object-field distinction and incorporate other concepts, for example the continuant-occurrent distinction by Galton (2008), to create a formal algebra of spatiotemporal data. The continuant-occurrent distinction is concerned with the identity of a phenomenon—a continuant is a geospatial objects that keeps its identity as changes happen to it, while the occurrent is in itself change, a distinct event that happens once. This conception moves many of Goodchild’s ideas closer to a practical implementation.

Ferreira et al. (2014) establish a clear split of spatiotemporal data into three types, by moving back to the principles of measurement: time series, trajectory, and coverage. Their model of how we observe this data is adapted from Sinton (1978). Sinton poses that to measure either space, time or theme (the domain of the data), we need to keep one of the three constant, change another one in a controlled way and then measure the third. Ferreira et al. (2014) reduce the six resulting possibilities to three by demonstrating how, in practice, three of the data types can be derived from the other three. The types of data are arranged as follows:

Time series Fixing space, controlling time, and measuring theme.

Trajectory Fixing theme, controlling time, and measuring space.

Coverage Fixing time, controlling space, and measuring theme.

Because these are realities of measurements, all geospatial data usually falls into these categories, even after further refinement or even if the data is fully simulated. Atop these directly measured data sets, Ferreira et al. (2014) build the derived data types of the coverage series (i.e. an ordered sequence of coverages such that changes in fields over time can be stored), the object and the event.

In Ferreira et al. (2014), these concepts appear as the component parts of an algebra of spatiotemporal data. This algebra is supposed to operate on a level of abstraction close enough to directly create GIS visualizations from it. As such, while it models the spatiotemporal realities of geospatial data in a usable way, its level of abstraction is lower than the one needed for our model. Simultaneously, Goodchild et al. (2007)'s conceptions are more removed from technical realities than would be helpful for the kind of taxonomy we want to create. We need a taxonomy of what kinds of GIS data are *possible*, not of what is currently feasible or mathematically definable.

With the literature lacking any established intermediate steps, what remains is to adapt these existing models to the needed level of detail—to merge the algebra defined in Ferreira et al. (2014) and the concepts from Goodchild et al. (2007), and to extend them as necessary into a taxonomy of possible geospatial representations, agnostic to most current conventions and technicalities of GIS. This taxonomy can then serve as a guideline of which kinds of representations can arrive in our embodied geosensification loops—a way to establish a taxonomy of the possible outputs of geospatial data sets that become the input for the sensory variable mappings from Section 3.1.

3.2.1 A GIS-agnostic Spatiotemporal Algebra

Ferreira et al. (2014) start at the following primitive data types, based on the OGC geometry model and the ISO temporal model: Value (Integer, Float, String, and Boolean), Time (Instant, Period), Geometry (Point, Line, Polygon, MultiPoint, MultiLineString, MultiPolygon).

First, the spatial data and value types have to be adjusted. The value types will be turned into the more abstract concept of scales, to fit in neatly with how the discipline of cartography looks at values (see Section 3.1). The geometry data types are instead replaced by Goodchild's concepts of geo-object and geo-field. Because they are most interesting for the power they have in defining the behaviour of values over space and the more unconstrained conception of object identity, we will treat them as purely geometric objects defined over three coordinates of space—the temporal axis will be handled by the two time data types. This constrains the space of possibilities in such a way that the resulting algebra retains a more realistic conception of what it is possible to represent. The data primitives thus are: Values (numerical, nominal, and ordinal), Time (Instant, Period) and Geometry (Geo-Object, Geo-Field).

After the data primitives are defined, Ferreira et al. (2014) go on to define their spatiotemporal data types. While redefining certain aspects of their algebra, we will stick to their notation style: The *type* notation defines the uppercase name of the type and lists its parameters and their types in square brackets. The *operations* are named in lowercase and denote functions. This is the syntactic part of the definitions, which defines names, domains and ranges. There is also a semantic part, which defines a set of axioms for these types. This latter part will not be included here, and instead only included in the full definitions in the appendix.

$$\begin{aligned}
&\textit{type } \textbf{TypeName } [ArgumentName : ArgumentType, ...] \\
&operations: \\
&\quad \text{operationname} : Domain \rightarrow Codomain \mid RangeCondition
\end{aligned} \tag{3.1}$$

We will keep Ferreira et al.'s basic building block of the *Observations* which represents a measurement of some fact which is locateable in space. This model however also applies to simulated processes as well as observation processes, as often geospatial data is the result of some computational model instead of a measurement. Simulations are thus considered to create artificial observations according to some computational model.

Observations are collections of tuples with three elements: a fixed attribute F, a positional attribute P, and a state-carrying attribute S. The fixed attribute can be a reference in space or time, or a set of reference values that identify a observation. The positional attribute is a collection of positions in either time or space, and the state-carrying attribute can be one or more thematic values or a (potentially moving) subset of space. In accordance with our goal for this algebra, what is shown in Definition 3.2 is a more abstracted version of the original definitions. It is however extended in one way: in this version, there can be multiple thematic values in one 3-tuple, as a simulation might yield a whole number of values in a single step. The following syntactic signatures however are still very direct adaptions from the original paper (Ferreira et al., 2014).

$$\begin{aligned}
&\textit{type } \textbf{Observations } [F : Type, P : Type, S : Type] \\
&operations: \\
&\quad \text{new} : \{(F, P, S)_1, (F, P, S)_2, ..., (F, P, S)_n\} \rightarrow Observations \mid n > 0 \\
&\quad \text{reference} : Observations \rightarrow F \\
&\quad \text{positions} : Observations \rightarrow \{P_1, ..., P_n\} \\
&\quad \text{sample} : Observations \times P \rightarrow S
\end{aligned} \tag{3.2}$$

Ferreira et al. (2014) resumes with an abstract definition for interpolators, as most observations will have to be interpolated in time and/or space in order to get a usable result from sparse measurements. This step will be skipped to stay at the desired level of abstraction. Instead we will resume right away with the abstract type *SpatioTemporal* that is defined as the basis for all the other spatiotemporal types.

$$
\begin{aligned}
&\textit{type}\ \textbf{SpatioTemporal}\\
&\textit{operations}:\\
&\quad \text{observations} : \textit{SpatioTemporal} \rightarrow \textit{Observations}\\
&\quad \text{begins, ends} : \textit{SpatioTemporal} \rightarrow \textit{Instant}\\
&\quad \text{boundary} : \textit{SpatioTemporal} \rightarrow \textit{Geo-Object}\\
&\quad \text{after, before, during} : \textit{SpatioTemporal} \times \textit{Time} \rightarrow \textit{SpatioTemporal}\\
&\quad \text{intersection, difference} : \textit{SpatioTemporal} \times \textit{Geo-Object} \rightarrow \{st_1, ..., st_n\} \mid\\
&\qquad st : \textit{SpatioTemporal}
\end{aligned}
\tag{3.3}
$$

For the intersection and difference, we assume that a geo-object is used as input. If necessary, a geo-field can be converted into a geo-object with a membership function, as explained earlier in this section.

From this abstract type, we can derive the actual spatiotemporal data types. Note that a coverage is not a pure geo-field, but a geo-field over which a boundary geo-object is defined. Its individual observations do not have to be points (like we are used to in current raster data sets), but are generally a statically defined collection of geo-objects over a geo-field.

$$
\begin{aligned}
&\textit{type}\ \textbf{TimeSeries}\ [G : \textit{Geo-Object}, T : \textit{Time}, V : \textit{Values}]\\
&\quad \text{inherits } \textbf{SpatioTemporal}\\
&\textit{operations}:\\
&\quad \text{new} : \textit{Period} \times \textit{Observations}[G, T, V] \rightarrow \textit{TimeSeries}\\
&\quad \text{values} : \textit{TimeSeries} \times T \rightarrow V\\
&\quad \text{min, max} : \textit{TimeSeries} \rightarrow V\\
&\quad \text{less, greater, equals} : \textit{TimeSeries} \times V \rightarrow \{ts_1, ..., ts_n\} | ts : \textit{TimeSeries}
\end{aligned}
\tag{3.4}
$$

$$
\begin{aligned}
&\textit{type}\ \textbf{Trajectory}\ [V : \textit{Values}, T : \textit{Time}, G : \textit{Geo-Object}]\\
&\quad \text{inherits } \textbf{SpatioTemporal}\\
&\textit{operations}:\\
&\quad \text{new} : \textit{Period} \times \textit{Observations}[V, T, G] \rightarrow \textit{Trajectory}\\
&\quad \text{geometry} : \textit{Trajectory} \times T \rightarrow G
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
&\textit{type}\ \textbf{Coverage}\ [T : \textit{Time}, G : \textit{Geo-Object}, V : \textit{Values}]\\
&\quad \text{inherits } \textbf{SpatioTemporal}\\
&\textit{operations}:\\
&\quad \text{new} : \textit{Geo-Field} \times \textit{Observations}[T, G, V] \rightarrow \textit{Coverage}\\
&\quad \text{values} : \textit{Coverage} \times \textit{Geo-Object} \rightarrow V\\
&\quad \text{min, max} : \textit{Coverage} \rightarrow V\\
&\quad \text{less, greater, equals} : \textit{Coverage} \times V \rightarrow \textit{Coverage}
\end{aligned}
\tag{3.6}
$$

It should be noted that the three *values* functions are not only not confined to returning a single thematic value or geometry, but are also not defined to take only a single point in time or space as an argument. They can be called with time intervals and many types of geometries and then return a wide array of thematically related values or geometries of the same identity. All that is required is that they all can be traced back to one reference in time, space or theme. In real-world applications this could for example mean that they are measured simultaneously, by the same sensor or that they together are the primary keys of a table in a relational database.

In order to allow coverages to change over time (measured for example by subsequent satellite flyovers), Ferreira et al. (2014) defines how to construct series of coverages.

$$\begin{aligned}
&\textit{type}\ \textbf{CoverageSeries}\ [G : \textit{Geo-Object}, T : \textit{Time}, CV : \textit{Coverage}] \\
&\qquad \text{inherits } \textbf{SpatioTemporal} \\
&\textit{operations}: \\
&\qquad \text{new} : \textit{Period} \times \textit{Phenomenon}[G, T, CV] \rightarrow \textit{CoverageSeries} \\
&\qquad \text{snapshot} : \textit{CoverageSeries} \times T \rightarrow CV \\
&\qquad \text{timeseries} : \textit{CoverageSeries} \times G \rightarrow \textit{TimeSeries}
\end{aligned} \tag{3.7}$$

These definitions can cover almost any sort of collected sensor data. But there are two issues: Trajectories do not contain values, while coverages do not have identity. This means that there is currently no notion of an object that can move through space while its internal values shift. This means—to go back to much earlier in this section—that there is still no object that can model a dynamic cloud in a GIS-focused way.

We thus go beyond the measurement data and construct real-world continuants (changing objects) and occurrents (one-time events) from the algebra, by merging the previously defined data types. The signatures of *Continuant* and *Occurrent* thus do not inherit from the *SpatioTemporal* abstract type anymore. These definitions are adapted from Ferreira et al.'s definitions of objects and events, however they change one important aspect: for objects it was assumed that each object only has a single changing attribute value over the whole geometry, presumably to keep their algebra true to what current GIS can comfortably model. Here, the definition is instead altered such that a continuant is a combination of a moving and shapeshifting, but identifiable geo-object, i.e. a Trajectory, and a CoverageSeries whose individual coverages only have non-null values wherever they intersect with the geo-object.

$$\begin{aligned}
&\textit{type}\ \textbf{Continuant}\ [ID : \textit{Values}, CS : \textit{CoverageSeries}, TJ : \textit{Trajectory}] \\
&\textit{operations}: \\
&\qquad \text{new} : ID \times TS \times TJ \rightarrow \textit{Continuant} \\
&\qquad \text{coverageseries} : \textit{Continuant} \rightarrow CS \\
&\qquad \text{trajectory} : \textit{Continuant} \rightarrow TJ \\
&\qquad \text{state} : \textit{Continuant} \times \textit{Time} \rightarrow (\textit{Geo-Object}, \textit{Coverage})
\end{aligned} \tag{3.8}$$

This, then, finally represents a hypothetical data type that would be able to fully model a dynamic cloud in a geospatial context. Almost any kind of phenomenon that can currently be represented in a real-time game engine but is too complex for GIS could be described by this data type: from animated objects to signed distance fields.

The occurrent on the other hand stays very close to Ferreira et al.'s original definition: It is defined by a geo-object that marks its area of relevance and carries its temporal duration, as well as a number of continuants that were involved in the event.

$$
\begin{aligned}
&\textit{type } \textbf{Occurrent } [ID : Values, G : Geo\text{-}Object] \\
&\textit{operations} : \\
&\qquad \text{new} : ID \times G \times \{con_1, con_2, ..., con_n\} \rightarrow Occurrent \mid \\
&\qquad\qquad con : Continuant \text{ and } n \geq 0 \\
&\qquad \text{time} : G \rightarrow Period \\
&\qquad \text{location} : Occurrent \rightarrow G \\
&\qquad \text{continuants} : Occurrent \rightarrow \{con_1, con_2, ..., con_n\}
\end{aligned}
\tag{3.9}
$$

Overall, this is a three-step process in which we start from a geographical observation, build varying kinds of spatiotemporal representation from it, and then define the concept of continuants and occurrents on top of these representations.

In order to move forward with the embodied geosensifications, we now need to start at this final step: how do we take the different types and make them comprehensible in an immersive environment? They need to fit together with concepts like sensory variables, which do not operate on the same level of representation as the continuant and occurrent. In fact, any system that makes use of the data types that are defined here will do so not on the types, but on the kinds of output they permit—down to whatever atomic types can actually be represented through a display system.

3.2.2 A Taxonomy of Spatiotemporal Output Types

In this section, we will move from data types to the "output types" of our type hierarchy and its operations. These output types are given by the totality of the operations in the algebra and are in a way a dissolution of the spatiotemporal data types into different spatiotemporal anchor points that can be sensorially represented. Instantiated in an actual application, these anchor points would appear as those interactive objects or scene elements that communicate something about a set of data, as opposed to any elements that are "set dressing" for a geographic scene or UI elements. These elements will always have a geometric component (because in this thesis we are only concerned with sensifying data that has a clear geospatial reference), and sometimes will have a temporal component and attribute values.

The operations as defined in the algebra are a valid way to construct such a taxonomy of output types, as they represent, as Ferreira et al. (2014) describes, a "typical" selection of minimal spatiotemporal and value operations that enable analytical work on changing data. Using the value domains these operations map to gives us a solid idea of the types of objects we can encounter for sensification scenarios, because more complex transformations are usually simply sequences of such minimal operations.

However, depending on the specific semantics of a function, a similar return type can have vastly different means in which it would have to be represented in a geosensification. For example, both CoverageSeries and TimeSeries have the *max* function that returns value representations. Seen in an interactive environment however, querying such features for their maximum values would have to be represented in very different ways: a CoverageSeries would show geometries in space in those areas in which its values are at a maximum, while a TimeSeries would show its geometry in space at the point in time at which its value is at a maximum.

To formalize this, we will consider the complete list of data types and their operation return values and attempt to establish a list of operations that yield semantically similar outputs. The results of this process are shown in Table 3.2. Some operations that are technically present in the definitions but do not have much meaning, like the *before* and *after* of one coverage, are not included. The *new* and *observations* operations are part of the functional part of a program (constructing objects and retrieving information from them) and are thus not explicitly listed for their output types. Any value-outputs, time-outputs and TimeSeries are given a geometry as far as it makes sense, as direct output of numbers, text, or graphs with a time axis into a scene is only possible by establishing some sort of UI space—an issue we will not deal with in this thesis.

We then remove all output types that map a data object to a spatially or temporally reduced version of itself. Now, if we combine the data types themselves as well as all the output semantics enabled by their operations into one list, we have a final taxonomy of possible geospatial output types. However, currently these types would all be very abstracted. In order to actually make statements about sensification or interaction processes, we need to establish one thing: their topological dimensions. As soon as we know whether a trajectory is a point or a line, we know in which ways it can change. A point can only move, but a line can also shift in shape. The same applies to the possible internal value distributions. Following Goodchild's conception of values in space, a point can only retain individual values, while a line can contain changes in these values over its (topologically 1-dimensional) geometric extent. Because we are in three-dimensional space, we can distinguish between points, lines, surfaces and solids. Solids will instead be called volumes, to better represent that in this kind of application they will usually contain volumetric fields of attribute values.

For attribute values, we already established the following types when we introduced the algebra: numerical, nominal, and ordinal. This mirrors the scale categories from Section 3.1, because these values are what will be represented by the sensory variables as the values defined over the spatiotemporal objects are mapped to them. For the purposes of this taxonomy however, it is only important whether values are present in a data type or not—conceptually, every spatiotemporal data type can be combined with every type of value and can carry as many values as necessary. One factor that is however important here, is the distribution of the values, i.e. whether the values are uniform over the whole feature (such as in the values that contain the identity of a geo-object) or whether they change throughout space. Thus there are three cases for each object: no attribute values, one or multiple value **Attributes**, or a **Field** of one or multiple values.

Data Type	Operations	Output Semantics
TimeSeries	boundary	**Bounds** of measured area
	less/greater/equals/after/before/during	Reduced *TimeSeries*
	values/min/max/begins/ends	**Values-at-Time**
	intersection/difference	Not applicable
Trajectory	begins/ends/geometry	**Shape-at-Time**
	boundary	**Bounds** over all movements
	after/before/during/intersection/difference	Reduced *Trajectory*
Coverage	begins/ends/after/before/during	Not applicable
	boundary	**Bounds** of Coverage
	intersection/difference/less/greater/equals/values	Reduced *Coverage*
	min/max	**Values-in-Space**
CoverageSeries	begins/ends/snapshot	**State** Coverage at time
	boundary	**Bounds** of a Coverage
	after/before/during/intersection/difference	Reduced *CoverageSeries*
	timeseries	**Sample** Timeseries over geometry
Continuant	trajectory	**Path** Trajectory over time
	coverageseries	The CoverageSeries **Basis**
	state	**State** Coverage at time
Occurrent	location/time	**Bounds** of event
	continuants	Continuant **Participants**

TABLE 3.2: The output semantics of each operation for each spatiotemporal data type. *new* and *observations* not included. Outputs that are spatially or temporally reduced but otherwise semantically identical to their input data type are marked in *italics*. Newly established output types are marked in **bold**.

Finally, there is the temporality of our objects. Here we can look back to Goodchild et al. (2007) and their characteristics of geo-objects over time. There is change of position, change of shape, and shape of internal structure. Change of position refers to the overall movement of the full object, change of shape refers to changes to the object's boundaries and topology, and the internal structure refers to the internal distribution of attribute values. Temporal channels thus have the following categorization of their possible dynamics:

Internal Structure **U**niform or **E**volving

Movement **S**tationary or **M**oving

Geometry **R**igid or **F**lexible

With this categorization we can describe the temporal characteristics with a simple triplet of letters. USR (Uniform, Stationary, Rigid) for example refers to an object that is unaffected by time throughout the whole data set, while EMF (Evolving, Moving, Flexible) refers to a highly dynamic object that evolves internally, moves through space, and changes its shape. This allows us to distinguish geo-objects into eight different temporality classes, in addition to the geometries.

With these three categories, we can fully describe the kind of output types that can arrive in our sensification. Instead of "moving cars as points", we have a collection of objects whose trajectories are topologically point-like, who over time remain uniform and rigid but are moving, and who contain attribute values like driving speed (numerical) and car model (nominal). This is a more complex definition than is usual for geovisualization, but as discussed, complexity is often needed to sufficiently model the dynamics of immersive geospatial scenes.

Considering all this, we can now categorize our output types based on a combination of the spatiotemporal data types, paired with a certain topology. This creates the final taxonomy that stands as the result of this section, and is shown in Table 3.3. This taxonomy will later integrate with the taxonomies in Section 3.1 and 3.3.

To explain three examples from the table, with the column in italics (topological dimension) and the row in bold (spatiotemporal data type):

1. A *Line* **Trajectory** is one in which a line-like object moves and changes its shape over time, but remains internally uniform in values (because trajectories are measured by keeping the thematic values static). The boundaries of its movement over time can be represented by a surface geometry in space. Its shape as queried at a certain time is a static line.

2. A *Surface* **CoverageSeries** is a geometry that contains changing values over the whole surface (i.e. it evolves internally). The outer bounds stay static throughout, because the individual coverages of such a series are all controlled to have the same geometry. It can be sampled at a certain point in time either with a sampling geometry (must be surface, line, or point), or by just returning the full coverage.

3. A *Volume* **Continuant** is a volume that is evolving internally, moving through space, and flexible in shape (EMF). Its values are fields that change throughout the whole volume. Its state can be sampled at a point in time, yielding a volumetric coverage. By outputting only the volumetric trajectory-part of the continuant, we can follow the path of the continuant without getting its values. The basis for its internal value-states is a volumetric coverage series that stretches over the whole extent of its movement over time.

Data Type	Data Dimensionality Point	Line	Surface	Volume
TimeSeries	Point ESR Attributes	Line ESR Attributes	Surface ESR Attributes	Volume ESR Attributes
Bounds	Point	Line	Surface	Volume
Values-at-Time	Point Attributes	Line Attributes	Surface Attributes	Volume Attributes
Trajectory	Point UMR	Line UMF	Surface UMF	Volume UMF
Bounds	Line	Surface	Volume	Volume
Shape-at-Time	Point	Line	Surface	Volume
Coverage	Point Attributes	Line Fields	Surface Fields	Volume Fields
Bounds	Point	Line	Surface	Surface
Values-in-Space	Point Attributes	Point Attributes	Point Attributes	Point Attributes
CoverageSeries	Point ESR Attributes	Line ESR Fields	Surface ESR Fields	Volume ESR Fields
Bounds	Point	Line	Surface	Volume
Sample	TimeSeries (Point)	TimeSeries (Line*)	TimeSeries (Surface*)	TimeSeries (Volume*)
State	Coverage (Point)	Coverage (Point)	Coverage (Surface)	Coverage (Volume)
Continuant	Point EMR Attributes	Line EMF Field	Surface EMF Fields	Volume EMF Fields
State	Coverage (Point)	Coverage (Line)	Coverage (Surface)	Coverage (Volume)
Path	Trajectory (Point)	Trajectory (Line)	Trajectory (Surface)	Trajectory (Volume)
Basis	CoverageS. (Line)	CoverageS. (Surface)	CoverageS. (Volume)	CoverageS. (Volume)
Occurrent				
Bounds	Point	Line	Surface	Volume
Participants	Continuant (Any)	Continuant (Any)	Continuant (Any)	Continuant (Any)

TABLE 3.3: The different spatiotemporal output types with their geometries, temporal dynamics, and possibility for values, given by the combination of spatiotemporal data types and their operations (rows) with the dimensionality of the data set (columns). Temporal dynamics and value possibility omitted where not applicable. "Attributes" refers to values that are uniform over the whole output geometry, "Field" refers to a distribution of values over the output. * denotes that the output geometry could also be of lower dimensionality, for example by sampling a point from a volumetric coverage series. (U = Uniform, E = Evolving, S = Static, M = Moving, R = Rigid, F = Flexible)

3.2.3 Conclusion

While the data types (TimeSeries, Trajectory, Coverage, CoverageSeries, Continuant, Occurrent) represent the data contents of our geographic scene, the output types will be the interface between this data and the rest of our sensification model. The basic idea behind these types is that the only way for us to sense or interact with geospatial data, is to represent it at a certain point in space, with a certain topological dimension. Once we can sense the geometric representation, it becomes possible to use our sensory variables to represent their values and to show changes over time.

This brings us to the last of our three taxonomies: Once we have perceived our data and its values through our senses, how do we interact with these geometries, so that we can move from a system that simply *displays* to a system that allows us to *analyze* data?

3.3 Interactivity—Embodied User Interfaces

By now, we have looked at taxonomies that categorize and conceptualize spatial data structures, as well as the human sensory system in the context of data sensification. What is missing now, is the glue that keeps them together—how do we select, modify and otherwise act on the spatial data so that it can reach our senses in ways that aid our sense-making? This, broadly, falls under the umbrella of interaction. It is only through interaction that we can start to interrogate complex spatial data structures. We want to enter a back-and-forth with the data and our own reasoning processes, in order to attain knowledge that we previously have not had.

How then can we classify and conceptualize the whole space of interactions? This is a complex area with a large number of taxonomies for different purposes and view-points. One can take a device-focused view, a body-focused view, a focus on a task that needs to be solved, and many more. One concept that appears again and again is the concept of "affordances", as defined in Chapter 2 of this thesis as the "possibilities of action that a certain technology, device or virtual tool enables for us", based on the definition adopted in Moloney et al. (2018). To summarize, the virtual environments we are looking at here must attempt to afford us possibilities of action that allow us to interrogate some spatial data that we perceive through several of our sensory channels at once.

Although there are multiple taxonomies for interaction design, we will stay the course of specifically working on concepts aimed at analyzing geographic information. Instead of the task-based approaches often taken in HCI research, which work well in the context of questionnaire-driven usability studies, we adopt the more general geovisualization-focused taxonomy created by Roth (2013b), which was already mentioned in Chapter 2. This taxonomy operates on a similar conceptual resolution as those developed in the last two sections.

In their taxonomy, Roth conceptualizes the process of interaction as the use of an *operator* on an *operand* to achieve an *objective*, which will aid in reaching a *goal*. They do this specifically in the context of geovisualization. In the course of their research, they identify several primitives for each of the four elements and explain how they can be made to relate. The goal of this section will be to see how the new interaction means embodied systems offer us would change parts of this taxonomy. We will start with the operands and the operators, with the latter requiring the most work, and then move on to the goals and objectives.

3.3.1 The Act of Interaction

First, there are the operand primitives, i.e. where in or on what part of a geovisualization an interaction happens. Roth (2013b) calls them "Space-Alone", "Attributes-in-Space", and "Space-in-Time". To simplify these operands in the larger system of taxonomies we are establishing, they will be defined as such:

1. Space: An interaction with just the geometry of the displayed data
2. Attributes: An interaction with the non-spatial values in the data that have been placed in space through the geometry
3. Time: An interaction with the time component of the data, potentially changing space and the attributes within

These primitives hold true for basically any visualization of spatiotemporal data, regardless of the environment or whether we are in an immersive system or not.

Only their relation and scale in reference to the user is changed, which is not a meaningful difference at this point in the taxonomy. To these operand primitives, we apply operators. The operators of a geovisualization are defined by Roth (2013b) as follows (with examples in brackets):

1. Enabling Operators:
    - Prepare for or clean up after work (import, export, save, edit, annotate)
2. Work Operators:
    - Changes to order and layout of the map (reexpress, arrange, sequence)
    - Changes to the design of the map (resymbolize, overlay, reproject)
    - Changes to the viewpoint onto the map (pan, zoom)
    - Examination of map features (filter, search, retrieve, calculate)

The enabling operators are basic computer operations, which will mostly stay the same in any kind of immersive system. The work operators on the other hand are the operators that are supposed to reach the desired objective, and are based heavily on the conventional GIS paradigms which inform most non-immersive geovisualizations. Some of them obviously do not translate well to the immersive case: *manipulating the user's viewpoint* is not necessarily something that is initiated as an operation from a UI interface—it simply happens as the user moves their physical body.

In fact, while not all operators become implicit in movements in this way, all operators in an immersive system have to be enacted *through* movements. While a non-immersive system would dress them as windows, context menus, keys and buttons, these concepts are inherently made for the mouse- and keyboard-based interaction paradigm. Here, interactions are triggered by pinpoint precision through small, 2D movements (mouse) or a large collection of keys that can be bound to tasks as necessary and almost always in reference to one 2D surface (the screen).

Interactions in spatial computing applications work through body movements, whether that be free movement, or movement that addresses worn, carried or otherwise tangible tools. The elements we are trying to interact with could be floating freely in 3D space or be bound to our body. Mouse and keyboard could be present as static tangible interfaces in an immersive environment, which we could optionally *move to* for specific tasks, but they have not gained much traction as the main control interface.

Keeping this paradigm shift in mind, we will now work through the work operator types. First, there are **operators that change the order and layout of the map**. The maps here would instead be the virtual immersive scene or the virtual scene elements. The issue is, that we will rarely display multiple scenes in the same way that we can display multiple maps. What we are more likely to change is the basic structure of the environment, physical laws like gravity or even flow of time, or our way of viewing a specific scene. This quickly bleeds together with **operators that change the design of the map**. For an immersive sensification this would describe how the environment is mapped to our sensory variables. Once we move away from layered maps and towards integrated virtual scenes, the distinction between these operator categories becomes blurry. This also includes **operators that change the viewpoint onto the map**. The equivalent of a zoom operation, i.e. rescaling the virtual camera that represent the users viewpoint, not only changes the viewpoint but also reshapes a part of the reality—the world becomes smaller. Similarly, changing the sensory mapping from a visual landscape to a fully sonic space completely alters any notion of viewpoint and arrangement.

The problems with the fourth kind of operator, i.e. **operators that examine map features**, are different. They might change certain parts of the environment in a limited way, but usually keep the world and reality intact. Here however, the act of examination explodes in complexity. Suddenly we can not only do operations like filter or search, we can move around to look at an object from all sides, we can touch and move and squeeze it, even taste it. The more dynamic and reactive objects become, the act of examining them could even veer into the territory of resymbolizing them—an object could for example start displaying an olfactory symbol whenever the user is examining it visually and haptically from close range (holding the object).

At this point it is necessary to reclassify the kinds of operators we want to distinguish. Rather than classifying them by type, we will classify them by how they "act on" parts of the sensification, similar to the three operand primitives. In this conception **enabling operators** remain intact as those operators that act on the application itself. Then we have **scene operators**, which act on the virtual scene, i.e. change the physics, appearance or arrangement of the virtual environment. Finally there are **sensification operators**, i.e. operators that act on the sensified data. These operators will be the most important for the majority of sensifications, as the sensified data is of course the main focus. Scene operators however can also be expected to frequently appear, especially in highly immersive non-situated contexts, where users can enact a lot of control on their surroundings. The most obvious invocation of such scene operators happens on the time operand, in which the current time is moved back or forth for all data sets in a virtual environment.

Another important aspect of operators in immersive environments is the associated body movement. While a pan or zoom operation remains similar for both mouse and touch interfaces, this is not the case in an embodied interface. Whether we resymbolize a feature by sniffing it (olfaction) or by throwing it into the air, where it creates a sound (sonification), is fundamentally important to the conception of the whole system. As such, we have to separate the concept of the operator itself, i.e. the action that is taken on the application, scene or data, from the act required to invoke it, i.e. the movement of the body.

For this, we can integrate the concept of *Embodiment* from Chapter 2. Operators in embodied geosensification are *embodied*. A filter operator could be embodied as a touch. Reprojecting the environment around us could be a spoken command. Resymbolizing a feature could be done implicitly by moving our nose close to it and smelling it.

Definition 4. ***Embodied Operators*** *are objective-driven user interactions on the application, scenes, or features in a sensification system, which are invoked through a specific movement or action of the user's body.*

Ideally, we could now specify the possible embodiments and operators as completely as we did with the sensory variables in Section 3.1. Roth (2013b)'s operator examples (import, export, save, edit, annotate, reexpress, arrange, sequence, resymbolize, overlay, reproject, pan, zoom, filter, search, retrieve, calculate) are a good start for the operator side. The list is not complete, but it clearly illustrates the kinds of operations that can be considered.

For the embodiment, we would need an equally complete list of possible body movements and actions. The issue is that the space of embodied interactions is a lot less constrained. We can go all the way from moving our body into a predefined area, to clearly defined hand gestures, to microgestural movements in individual finger joints (Dementyev and Paradiso, 2014; Berger, 2021a), or even to systems that enable interactions like speech recognition of mentally verbalized but unspoken words by measuring otherwise imperceptible, subconscious neuromuscular movements (Kapur et al., 2018). And even when staying at a certain level of bodily resolution, the possibilities of action are endless. Examples from the literature feature embodiments such as squeeze, crunch whole, crunch partly, nip, rip (a piece of paper in Engeln et al. (2018)), grabbing, shaking, throwing (a virtual map in Newbury et al. (2021)) or even eating (a cake representing data in Mueller et al. (2021)).

In order to establish a taxonomy, what we need instead is a clear conception of how to formulate the idea of an embodied operator. What we *can* usually specify about the embodiment, is which body part is involved in a broad sense. Whether we use a grabbing gesture or subconscious finger twitches, both embodiments broadly originate from the hand. This can be especially important in combination with the sensory variables from Section 3.1, as there can be obvious synergies (like a grasping gesture triggering a haptification in the hand) as well as antagonistic effects (like requiring the user to speak to activate a gustation display that is already placed on the tongue). However, we do not always interact with just a body part—frequently, we are wielding some sort of tool that allows us new possibilities of action. For the purposes of this taxonomy, we will treat body parts and tools as one and the same (further discussion on this will also be part of Chapter 4). To name this combination we will adapt a term from the field of robotics and refer to body parts and tools broadly as **effectors**.

Given the operator and its effector, what is missing is a way to formulate the specific action that is supposed to be taken. One thing one will notice when considering embodiments in the context of user interaction, is that their explanation frequently falls back on a verb (see examples such as “crunch” or “eat” earlier in the section). What really differentiates the taxonomy presented here from the one presented in Roth (2013b), are the verbs that we use to apply the operators. Our body parts or tools simply allow us to enact these verbs, i.e. they *afford* us certain verbs that we can then link to operators. Surprisingly there is a clear overlap here with a discipline that has influenced scientific research into immersive experiences more than perhaps any other factor over the last decade: game design. One of the main reasons to even develop an embodied sensification system is to allow users to engage with the data—and what is more engaging than products specifically engineered for engagement? One of the most important concepts in game design methodology are the concepts of *Verbs* and *Objects*. Verbs are what the rules of the game permit us to do to the objects in the game world. The similarity to our embodied geosensifications should be obvious: Objects are our geospatial features, more specifically their operand primitives. Verbs can be seen as the operators that our video game avatar can enact on these operand primitives. And because our sensifications are embodied, we ourselves of course take the place of the video game avatar.

By adopting this approach, we arrive at the following: an embodied operator is defined through an **Effector** enacting a **Verb** to trigger an **Operator**. Figure 3.3 shows this concept in reference to the user and the operands that the sensification of a spatiotemporal data set enables.

Roth (2013b) includes several tables containing examples of certain combination of parts of their taxonomy. These largely remain intact after our changes, though the scope and use of the examples would change to something more appropriate for an embodied sensification. To complete our change to the taxonomy, Table 3.4 shows

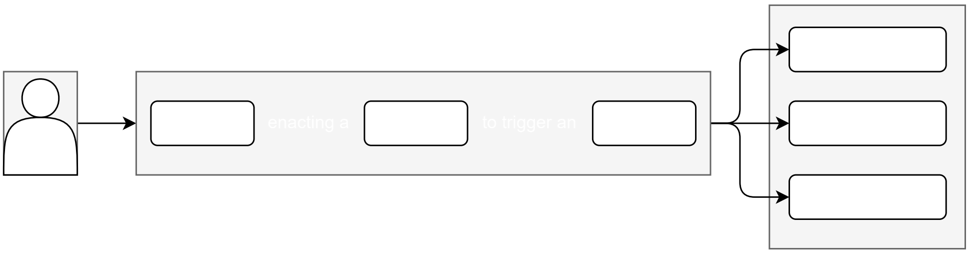


FIGURE 3.3: Concept of a user interacting with a data set in an embodied geosensification by using an embodied operator on the data's operands.

Operator Type	Effector	Verb	Operator
Enabling	Mouth	**Speaking** data set name	Import
Enabling	Arm	**Throwing** feature upwards	Export
Scene	Arms	**Running** hand **over** forearm	Alter Time
Scene	Hands	**Moving** hands **closer** and **apart**	Scale Environment
Scene	Leg	**Stomping** on the ground	Alter Environment
Sensification	Mouth	**Blowing** air onto a feature	Filter Feature
Sensification	Nose	**Inhaling** close to a feature	Retrieve Value
Sensification	Lens	**Gazing** through	Search Features
Sensification	Wand	**Touching** two objects	Calculate Distance

TABLE 3.4: A list of examples for operator embodiments, consisting of effector, verb, and operator, sorted by the operator type. The lens effector refers to a virtual lens that can be picked up and moved around, the wand effector refers to a virtual handheld stick, possibly lining up with a haptic controller the user is holding.

examples of embodied operators. It should be noted that the concept of verb does not require its content to be just one word.

3.3.2 The Purpose of Interaction

The first step to describing *why* an interaction takes place are the objective primitives as originally stated by Roth (2013b), in rising complexity:

1. Identify: Retrieving information about a specific map feature.
2. Compare: Finding similarities and differences between map features.
3. Rank: Establishing an order for a collection of map features.
4. Associate: Characterizing correlations and dependencies in a collection of map features.
5. Delineate: Organizing map features into categories and clusters.

Each objective is usually the result of one embodied operator being invoked on a set of data, i.e. one single act of interaction. Once we fulfill a certain number of objectives by applying operators to operands, we hope to reach a goal. Reaching a goal means that we attained some knowledge or wisdom based on our interaction with the scene or data. Roth et al.'s taxonomy distinguishes between three different levels of goal, also in rising complexity:

1. Procure: Retrieving information.
2. Predict: Creating a forecast.
3. Prescribe: Aiding in decision-making.

The set of objectives and goals does not meaningfully have to change for the embodied case—in the end, we are still trying to make sense of a collection of data, and these objectives and goals are the basics of geospatial sense-making. However, this work is not just about embodied interactions—it is also about working with geospatial data in immersive and situated geographic scenes. Especially for the situated case, we might not only be interested in these five conventional data analytics steps, but in much more integrated objectives. Let us imagine for example an ecological experiment which features researchers in the field choosing in which location to work based on data supplied by a situated sensor network. Here, interactions could have objectives like *locating* a feature after one sensor reading triggered a warning.

These more environmental interaction objectives need to be part of our taxonomy. One way to think about this problem, which will also become relevant again in Chapter 4, is to think about embodied user interfaces not as analytical interfaces, but as *powers* that the systems gives us in relation to an environment. This follows a conception taken in Willett et al. (2021), where superpowers from popular culture are used as a framework to inspire visualization techniques. The authors focus on powers that allow characters to "gain knowledge about things, people, or phenomena" through visual means, and establish a taxonomy of seven *enhancements* these powers tend to give their users: enhanced vision, visual synesthesia, enhanced attention, enhanced comparison, enhanced numeracy, enhanced prediction, and enhanced recall. Unsurprisingly, some of the powers shown in the paper also have a geospatial focus. Visual synesthesia (i.e. the translation of a non-visual phenomenon into a visual representation) for example is explicitly explained through the example of a situated geovisualization: White and Feiner (2009)'s SiteLens air quality measurement visualization system, in which air quality measurements are visualized in-place through an AR display. Figure 3.4 shows this and other examples of such powers.

Almost every sensification is going to involve some form of such a synesthesia—in a way it is the purpose of the whole concept of sensification, to show a sense something it can not normally grasp. Similarly, enhanced vision would have to be expanded to all the senses. In fact, for the purposes of sensification, their separation quickly becomes meaningless: both of these enhancements in the end just **enhance perception**. In this way, they can be the objective of an interaction. An interaction could activate some sort of new sensory mapping, like moving a virtual lens in front of our field of vision to gain a form of geospatial "x-ray" vision for everything behind the lens, as for example in Kluge et al. (2019).

The other enhancements can also be seen as interaction objectives. **Enhanced attention** is about better recognizing small or important details in the environment, and could thus cover the *locate* objective we named earlier. Here we have to discern them clearly from the *identify* goal, which is focussed on closer examinations of individual features rather than perceiving them in the first place. **Enhanced comparison** maps directly with three of the existing objectives: compare, rank and associate.

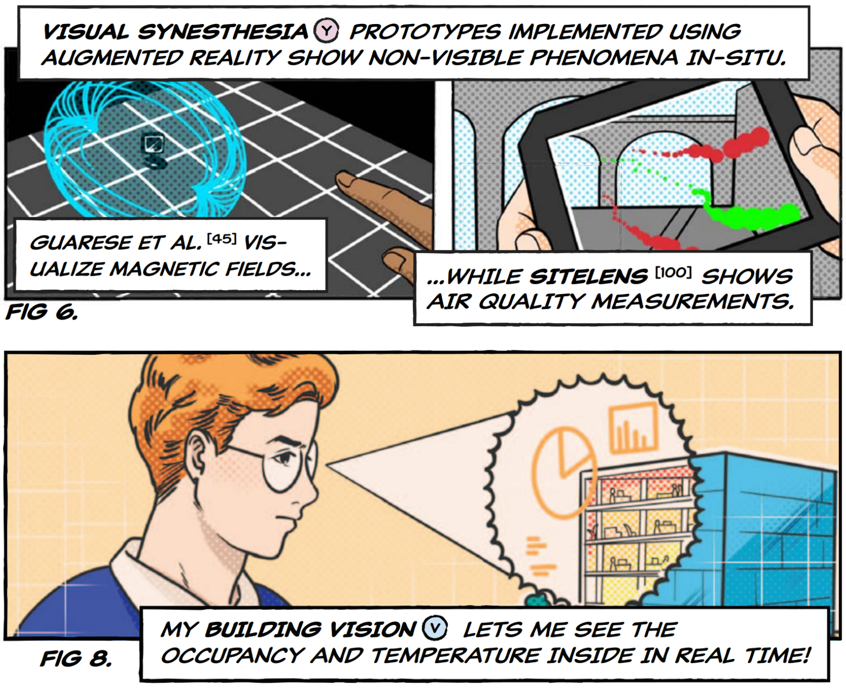


FIGURE 3.4: Two examples of situated geospatial visualizations conceptualized as superpowers, combined from Willett et al. (2021).

Enhanced numeracy is about counting, measuring and aggregating information. It has some overlap with the existing objectives of associate and delineate, however its main aspect of counting and measuring are novel in relation to the objectives. An example for enhanced numeracy would be a system that assists a drone pilot in situated awareness of how close their drone is to both the areas of interest as well as potentially dangerous obstacles.

Enhanced prediction appears as a goal in Roth (2013b), however in the way Willett et al. (2021) conceive of it it is *also* an objective: a situated system could for example show what the weather will be in several hours, in which case prediction would be a goal, but a system could also let the user query a short prediction of the movement of a ship on a shipping route, in order to better fulfill the goal of *procuring* information about the ship, like its course or speed. This then leaves **enhanced recall**, which maps to the objective of association, but does so over time rather than over space.

Because there are both overlap as well as differences between Roth (2013b)'s sense-making objectives and Willett et al. (2021)'s perceptual and cognitive enhancements, we can combine them into one taxonomy of *enhancements* that covers the typical applications of both situated as well as fully immersive interaction scenarios. As a naming scheme we will stick with verbs as first established by Roth (2013b). The following list contains the names and an explanation that finishes the sentence: "The interaction has the objective of enabling us to...":

Observe ... perceive features in a new way.

Notice ... recognize important details about features.

Examine ... identify information about a feature.

Compare ... compare two or more features.

Rank ... determine the order of multiple features.

Associate ... determine relationships between multiple features.

Enumerate ... count or aggregate specific facts over multiple features.

Measure ... measure or estimate facts about one or multiple features.

Predict ... make prediction about the movement or evolution of features.

Delineate ... organize features into logical structures.

Considering these enhancements, what changes about our goals? Every enhancement is in service of at least one of the goals. Examining and comparing a feature are in service of procuring information about the larger data set. Delineating large amounts of features shows trends and can thus be used to predict and to prescribe what should be done. As such, we can look at the newly added enhancements and see whether they require new goals as well, or whether they are already working towards the three existing ones. The newly added enhancements are *observe, notice, enumerate, measure* and *predict*. The first two imply a passive awareness about the immediate surroundings, which is not about procuring specific information. This goal could thus be called "Perceive". Enumerating, measuring and predicting still fall under the original three goals.

It is important to note however, that every objective can be in service of every goal. There are some general tendencies of objectives that will appear often when trying to attain a certain goal, shown in Table 3.5, but an example could be constructed for every combination.

Goal	Enhancements
Perceive	Observe, Notice, Examine
Procure	Examine, Compare, Rank, Enumerate, Measure
Predict	Rank, Associate, Measure, Predict
Prescribe	Associate, Predict, Delineate

TABLE 3.5: A table associating the four interaction goals of our taxonomy with the most commonly employed cognitive enhancements needed to reach them.

3.3.3 Conclusion

Once we reflect on the concept of enhancements, we will notice something: they are triggered by an interaction, but they will usually alter how the data is displayed. We touch an object to examine its attribute, and the reaction of our application is to show these values in various ways. In such cases the goals are actually reached by the sensification, not by the interaction—the interaction just prompts the sensification configuration to change such that we can derive the desired information. It should be very apparent now that these three created taxonomies must be integrated with each other to allow us to conceptualize systems in their totality. So what is the purpose of interactivity in relation to the whole system?

Interaction is the glue between the user's senses and the displayed data. Without it, there would be no sense-making beyond whatever the original creator of a static image or a recurring sound intended to show. Interaction turns a multi-sensory immersive display from something that is showing a collection of facts to a user, into an embodied analytics interface. The interactivity is our main interface to everything else that has been established in this thesis. As such, they are perhaps the most important but complex aspect to consider in building such applications.

If we once again take a step into the realm of game design, there is a common wisdom that what makes a game fun and engaging is what arises from the combination of all the verbs it contains. This carries over to embodied geosensifications—there is hardly a reason to go through all the implementation and hardware trouble of current spatial computing systems, only to create a sensification that only has verbs that are similar to the ones that any desktop visualization would have. Another important aspect of game design is that it is often better to have very few verbs but to make them as meaningful and engaging as is possible. This holds true even for scientific interfaces—interactions do not always need to be *fun*, but they need to be easy to learn and understand, and reactive and engaging to use, especially when asking users to accept and learn entirely novel interface paradigms. Clarity about data can be achieved on a monitor—we instead have a possibility here to foster different kinds of understanding through meaningful, layered engagement within data-driven virtual environments.

With these principles in mind, we can now move on to Chapter 4, in which the taxonomies will be combined and expanded into a full model and diagram language. At the end of this process we will also return to the superpower-based conception first introduced in this chapter, in order to create the model in such a way that it promotes establishing a limited but strong set of verbs.

Before this, the results of the current chapter can be summarized in one sentence: We try to establish a small set of meaningful and evocative embodied operators, that we can apply to the operand primitives of our sensification, in order to enhance certain kinds of cognitive processes within the virtual scene, ultimately causing us to reach goals in understanding one or multiple data sets.

Chapter 4

A Model for Embodied Geo-Sensification

4.1 The Original Model

Now that we are equipped with conceptual tools and taxonomies for several different aspects of geospatial sensifications, there are two questions: How does it all fit together into a cohesive whole? And how do we put this cohesive whole into practice?

To start, we will take a step back—to an early draft of the thinking that led to what was shown so far in this thesis. Specifically, a proposed new kind of diagram specifically focused on modelling geospatial sensifications. These ideas were published in Berger (2021b) and shown at the 12th International Conference on the Theory and Application of Diagrams in 2021. Parts of the contribution will be included in the following explanations.

We begin by asserting that the diagrams we use to model complex geospatial data visualization are often based on common diagrams from computer science, like the suite of UML diagrams. However, "these diagrams rarely explicate the larger concept of the modelled system, instead focusing on specific programming choices and subsystems. This can become an issue for visualization systems that do not make use of established methodologies, as the very technical diagrams might mask some of the novelty." (Berger, 2021b)

This masking is an effect of the chosen layer of abstraction. Many established diagrams highlight very specific aspects of a system, like a sequence of networked communications, or a collection of use cases for different types of user. What is more important than anything else in a visualization systems however, is its core sense-making loop. This loop heavily integrates with the given interaction devices and display system. The interrelations at this level of abstraction are important enough to deserve their own diagram language.

This is where our taxonomies come into play. They are taxonomies describing different aspects of embodied geosensifications—of systems that place a user "into a virtual environment that *is* the data, which they then experience through multiple senses, while simultaneously manipulating it with the same body that is also "carrying" the aforementioned senses. In this way, the body is made the center of a multivariate feedback loop." (Berger, 2021b)

The concept of this feedback loop is crucial to embodied sensifications. And this should not be surprising—in the state of the art, it was already discussed how modern conceptions of visualizations and the visualization pipeline show a visualization feedback loop instead of a straight pipeline. For this reason, our attempt to model embodied geosensifications diagrammatically starts with McCormack et al. (2018)'s multisensory analytics loop. Here, "they extend the traditionally visual mapping step to a *sensorial mapping*, which is defined as 'a mapping from data elements and data attributes to sensory channels (sight, hearing, touch, proprioception, smell and taste) and their respective sensorial variables (color, pitch, roughness, etc.)'. The rendering stage is extended into a device-focused stage, where a device enacts every one of the sensory channels onto the human body, and a body-focused stage, where the involved human senses perceive and body movements trigger interactions."(Berger, 2021b)

To simplify, we use the concept of embodiment to close the separation between body and display devices and to combine them into one. "There is no way to present a digital sensory channel to a body part if there is no device that can do so" (Berger, 2021b). We also extend the interaction steps of the pipeline, to make it clear that interactions can change the original data as well as the specific variable mapping. The result is shown in Figure 4.1.

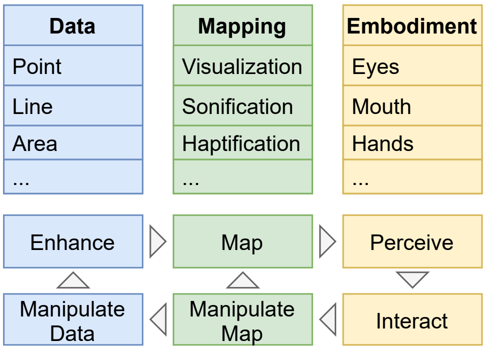


FIGURE 4.1: The multisensory sensification and interaction loop, adapted from McCormack et al. (2018). From Berger (2021b). Area data refers to surface data as defined in Section 3.2.

To go from this conception to a diagrammatic syntax that can be more specific about the connections between the different steps, we go on to introduce the UML-adjacent concept of message-lines, which highlight the feedback loops between sensory mapping (data to body) and interactions (body to data). The concept of spatial data types, attribute values and spatial transformations (which are often necessary to go from a raw spatiotemporal data set to whatever can be usefully rendered to a display in response to some interaction) are also introduced. The way this all works together in Berger (2021b), is shown in an abstracted sample graph in Figure 4.2.

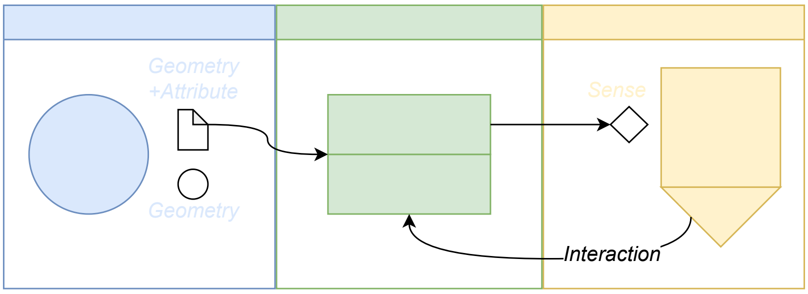


FIGURE 4.2: The sensification loop from Figure 4.1, conceptualized as a diagram. It shows how the feedback loop is created by transforming both geometry and attribute values of a piece of data into something that can be sensed and interacted with by the user's body. The body is shown through body parts, which can have one or multiple senses and one or multiple movements (interactions). Three swimlane boxes separate the steps of the diagram into data, mapping, and embodiment, to highlight the relationship with the sensification loop. Simplified from Berger (2021b).

Additionally, four different types of message lines are introduced, highlighting the timings by which the sensory presentation of a piece of data can arrive at the user: "queried (message is sent once upon an interaction), feedback (message is sent continually as interaction happens), interrupt (message is sent once a specific state is reached) and continuous (the message can change independently of interactions and is sent constantly)." (Berger, 2021b)

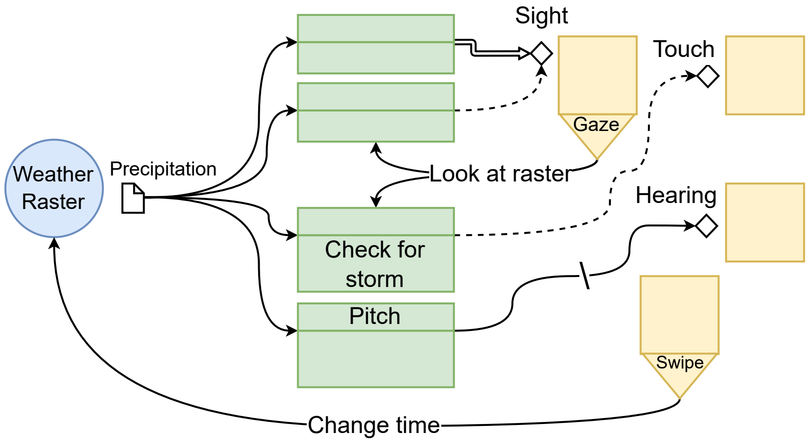


FIGURE 4.3: An example diagram for an immersive weather sensification, based on Figure 4.2. The beaufort scale is a scale for wind forces. Note that this example diagram does not include the swimlane boxes and instead communicates the type of element by color: blue represents data elements, green mapping elements and yellow embodiment elements. A dotted line represents feedback, a crossed line an interrupt, and the double line a continuous message. Lines travelling backwards denote interactions. Simplified from Berger (2021b).

With this basis, a real world use case was considered: "We imagine a user wearing a pair of augmented reality glasses with hand interaction and vibration functionality. Running on the glasses is an augmented reality weather system. Users would be walking around in the real world and could see weather maps directly projected into the sky. An approaching rainstorm would regularly emit pitched warning sounds as it approaches. If a user looks up, more detailed textual information about a cloud could be displayed, and a rumble in the AR glasses could signify that a cloud the user is currently looking at is a storm center. With a hand swipe the user can fast forward through the predicted weather." (Berger, 2021b) The result of translating this use case into a diagram is shown in Figure 4.3.

This diagram was the result of Berger (2021b), which concluded by stating that "the diagram can model a system with several modalities and interaction types, and it can promote thinking about specific steps and necessary transformations." As such, this diagram could serve as the base for a diagram language that goes along with the model that we will create throughout this section. In this original form it was however not backed by the extensive review and taxonomies that we established over the last two chapters. As such, changes and extensions are necessary to turn this basic diagram into the appropriate tool that brings together the conceptual foundations that have been laid so far.

To structure this process, we will establish a list of necessary changes and extensions, based both on feedback collected about the original publication and the new lessons learned in this thesis. We will then go through the points one by one, constructing an appropriate model and diagrammatic representation. Where necessary we will trace the changes through the weather sensification example from Figure 4.3.

The list consists of ten points and goes as follows: (1) The behaviour of geometry-only data in which there are no attributes to be sensified needs to be fully included. (2) The concept of data temporality needs to be included. (3) Direct manipulations on the data as shown with the "Change Time" interaction in Figure 4.3 need to be fully conceptualized—this currently appears as a change to the data, which is not a common occurrence in actual visualization systems, as feeding changes back into the original data is a process entirely separate from the act of visualization. (4) The role of the virtual or situated environment for the model needs to be clearly established. (5) Tools and tangible objects need to be integrated. (6) The diagram should allow us to highlight to the reader what the purpose of each feedback loop as well as the whole system is. (7) The combination of senses, space and time needs to be properly conceptualized, to cover constraints like haptics only being able to interact with data at short range or vision not being able to see through surfaces. (8) The graph needs to be able to scale well over varying system complexities—it needs to be appropriate both for mono-sensory sensifications with one interaction, as well as fully integrated sensification systems with many moving parts. (9) Message types need to be conceptually integrated with the three taxonomies. (10) There should be a clear list for the scope of each element in the diagram, and a way to construct a diagram step-by-step.

Problems (1) through (7) will be solved one-by-one throughout Section 4.2 by applying and adapting the three taxonomies from Chapter 3. Problem (8) through (10) will be solved in one larger step (Section 4.3), by going back to superpowers-framework first mentioned in 3.3 and using it to further integrate the established concepts into a diagram language that can not only be put to the page, but can be used as a template to construct embodied geosensifications based on common metaphors and step-by-step processes.

4.2 Applying the Taxonomies

With a list of possible extensions and changes to the as-published model established, we can go through them one by one and see if the sense, data and interaction taxonomies can apply to them and help focus in on a solutions.

4.2.1 Decoupling Geometry and Values

First, we need to acknowledge that we can not treat every act of data representation as going through some sensory variable. The purpose of the sensory variables is to fulfill an act of data representation that is numerical, ordinal, nominal, selective or associative. Most of the literature does not make this distinction explicit, but every time we only display the raw geometry of a data set, we might be using multisensory display technologies, but we are not encoding any attribute values into this data stream. Showing the true geometry of a polygon is an act of establishing a spatial reference, which is a different act than representing a data attribute. Geometry has a special role in the realm of geospatial data.

How can this be represented more fittingly in the diagram? First we can acknowledge that the act of displaying geometry can also be made to different senses, just like the act of sensory mapping. The "Spatial Transform" step in the original diagram can actually be conceived as the mapping step for exactly this geometry-display. It takes the data geometry, and turns it into something a sense can perceive (like polygons being transformed into centroid points for sonification). A clear definition of this will follow after the concept is extended in the next section.

In fact, a sensory mapping is not possible without some sort of reference geometry that is displayed first—we would have no concept of which geospatial feature we are learning information about. For the diagram, this means that we need to separate the connectors for geometry and value, and also separate the sensory and spatial mapping. What applying this new conception to the abstract diagram from Figure 4.2 could look like is shown in Figure 4.4.

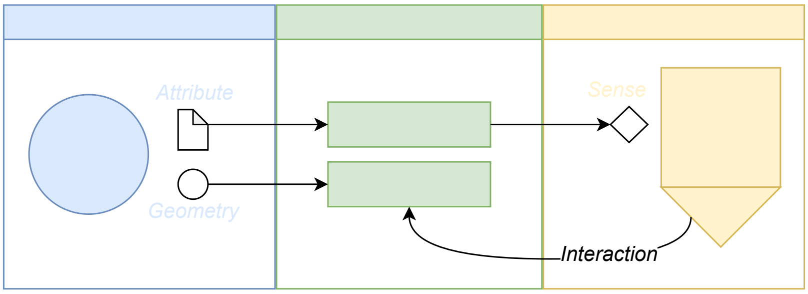


FIGURE 4.4: Changing the diagram from Figure 4.2 to distinguish sensory mapping of values and spatial mapping of geometry. "Sensory Variable" and "Spatial Transform" are separated, and "Geometry+Attribute" was changed to just "Attribute".

One aspect to note for this new syntax, is that the spatial reference a sensory mapping is operating on is implicit. There is no direct line between the "Sensory Variable" and "Spatial transform" box, because as soon as the geometry of a data set is displayed in some way, we have a reference that values can be displayed in accordance to. Sometimes this is direct, like visualizing a color on a visualized geometry, but it can also be prompted by interactions, like playing a vibration once we point a handheld controller at the geometry.

4.2.2 Temporality of Data

As of Section 3.2, the temporality is part of our conception of data. Currently the diagram is only concerned with geometry, how it is changed during the mapping stage (e.g. creating a buffer around points so that they are circles that can be visually perceived), and how it appears to the observer.

According to our taxonomy of output types (see Table 3.3), there are three possible aspects to a data set that can be brought into the sensification loop:

Geometry One of: Point, Line, Surface, Volume

Temporal Dynamics One of each distinction: (U)niform or (E)volving, (S)tationary or (M)oving, (R)igid or (F)lexible

Attribute Values One or multiple values, each with a type according to Section 3.1, that are either attributes or fields.

For each type of geometry and temporality, there is also a finite list of sensory modalities we can even display to, just like each type of value has a finite list of useful sensory variables that it can be displayed through. This gives the diagram more predictive power, as design possibilities are constrained to some degree. This list of mappings will be made explicit later in this chapter in Table 4.2. For now, lets look at Figure 4.5 to see how our diagram changes according to the new conception of spatiotemporal outputs. Geometry, temporal dynamics and attribute values are now separated into individual stacked rectangles. The input and output points are not separate symbols anymore, as writing their meaning out as a word is easier for a reader to understand than having to learn a collection of different symbols.

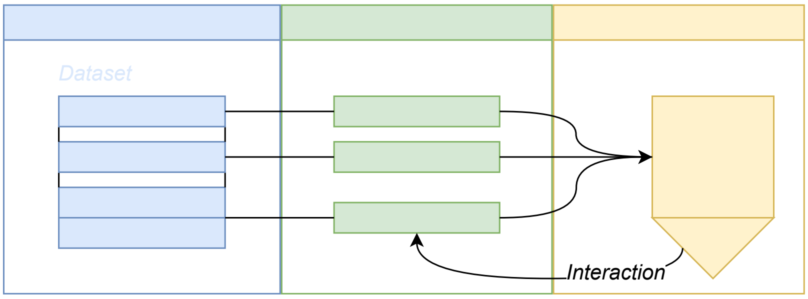


FIGURE 4.5: Changing the diagram from Figure 4.4 to include the three different kinds of output each spatiotemporal data set has. The graphical input and output nodes were removed in favor of direct connections between the data, mapping and embodiment elements. This abstract example assumes a data set with two different values.

What was originally a "Spatial Transform" is now simply called a "Transform". More technically, the concept is now changed into that of the "Spatiotemporal Transform" (see Definition 5). This transform can either attach to the geometry or temporality of a data set (or both).

Definition 5. ***Spatiotemporal transforms*** *or* ***spatiotemporal mappings*** *are any sort of GIS operation that change the geometric or temporal representation of a data set while keeping the attached attribute values intact, in order to display the data to one of the user's senses in a way that is appropriate for the current sensification context.*

To illustrate what the changes so far mean in practice, Figure 4.6 shows an adaption of the situated weather sensification from Figure 4.3. We can see which sensifications happen on which attribute, what the scales of the attributes are, and which mappings represent spatial and temporal mappings.

However, even though the diagram is more precise in many aspects, there are some new issues that need to be resolved: it is unclear which sensory mappings happen in reference to which interaction, and the different types of mapping appear exactly the same on first glance, overcrowding the diagram. This will be solved later on in this chapter. As we will see, the issue lies in the lack of separation between the part of the example that informs the user about precipitation, and the one warning the user about storm formations.

4.2.3 Interaction Elements

An important concept in Berger (2021b) was that interactions would either be done to change the underlying data, or they would act on the mapping of a visualization. In the weather sensification example, this is shown by the difference between the "Look at raster" interaction, which retrieves a value, and the "Change time" interaction, which moves the current time of the weather data back and forth (historical weather to predicted weather).

While this was in some ways a useful distinction then, it implies something that is not true: That we are changing the same data that we are sensifying. This is not what generally happens in real applications. Sometimes we use sensifications to figure out missing pieces or errors in a data set, and then go on to fix them by filling in data, as seen with the *enabling operators* from Section 3.3. However, this is not

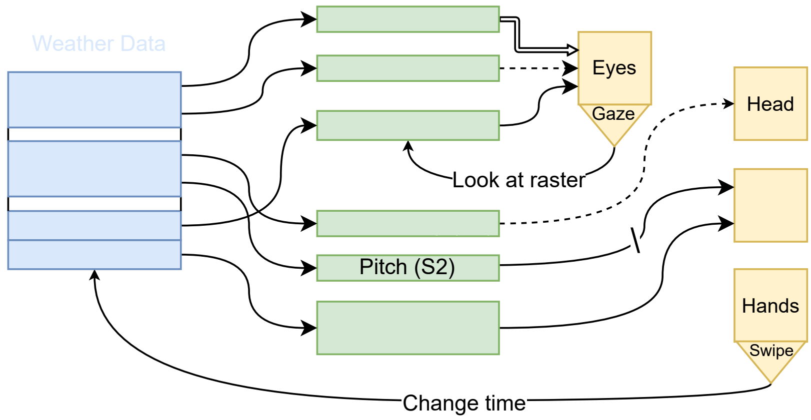


FIGURE 4.6: The example from Figure 4.3, changed in accordance with new concepts introduced in Sections 4.2.1 and 4.2.2. Precipitation and Beaufort scale are the two values in a weather raster. The weather raster is represented as a surface, whose internal values evolve over time (ESR—Evolving, Static, Rigid). Blue represents data elements, green mapping elements and yellow embodiment elements. A dotted line represents feedback, a crossed line an interrupt, and the double line a continuous message. Lines travelling backwards denote interactions.

strictly part of the sensification feedback loop as we have conceptualized it so far, in which we react to sensory inputs with interactions, which then change the sensory inputs.

In these feedback loops, we almost always act to change the mapping. In the original diagram, this is shown by most of the interaction lines going back into the spatial transform part of the mapping. This however was the main cause for confusion when presenting the diagram to others. Let us take a look at Figure 4.3. We are "Gazing/Looking at a raster" but the interaction line runs to a sampled point. The raster is invoked in the label on the interaction line, but the line ends on a point. This is an inherent disconnect in conceptualizing interactions as something that is happening on *a mapping*. For the loop, it is much more meaningful to think of interactions as something happening on *the data*, which then results in changes to the mapping. In this example, we are looking at the raster, i.e. the geometry of the basic data, and then the value at the point we are looking at is sampled (spatial transform), and displayed as a text label (sensory mapping).

What solves this issue, is the fact that the taxonomy in Section 3.2 does not describe sets of data, but those parts of the geospatial data that can be introduced into a sensification (the output types). This allows us to properly target interactions at the data, without implying that the original data is being changed—we are only changing the output of an operation that was called on the original data (either one that directly returns a full representation, or the algebra operations like *boundary* or *sample* defined in Section 3.2). This is best represented in the diagram by including a reference to the raw spatiotemporal data set as something external to the diagram, while the diagram only represents one specific sensification feedback loop, in which interactions run from the body to a certain aspect of the data.

This fits directly into several of the concepts from Section 3.3. Here, embodied operators are used to act on one of three operand primitives. These operand primitives cover interactions on the geometry (*Space*), interactions on the sensory mappings (*Attributes*) and interactions on the temporal aspects of the data (*Time*). The operands correspond directly with the three output channels of our diagram (geometry, spatiality, attributes). The embodiment of the embodied operator is already given by the body part (effector) and its interaction movement (from here on called *verb*, as in Section 3.3)—only the operator itself would have to be added.

A possible way to include this in the diagram is shown in Figure 4.7. The interaction is added as a new type of element in the color red. At this point, the movement is also placed into the same box and covered by the textual description, in order to not overcrowd body parts with diagram elements just because they can perform many different types of movements, like the hand in a gesture-based system.

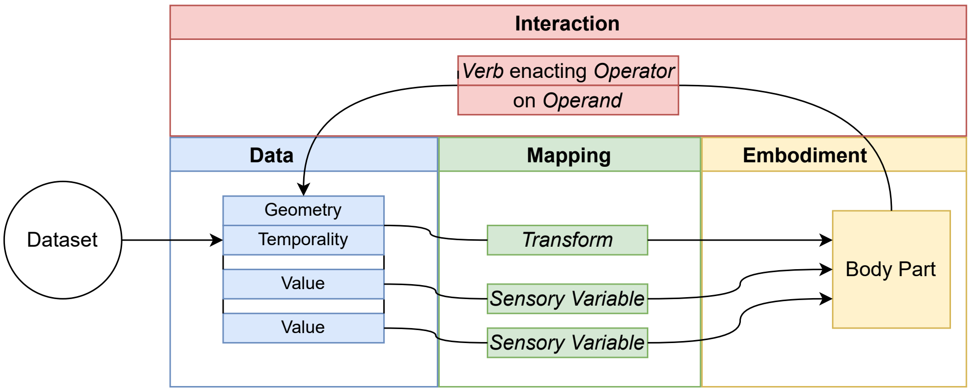


FIGURE 4.7: An alteration to the diagram from Figure 4.5 that makes all interactions act on the data step of sensification, but also conceptualizes the sensified data as a output/subset of some original raw data. The concepts of operands and embodied operators from Section 3.3 are now also included as an aspect of the diagram, within the new *interaction* category.

Treating the data as an outside factor that only supplies the output channels (and thus the interaction operands) also introduces another advantage: We can now define aspects outside of the data-embodiment loop. Another outside aspect that is also relevant for the whole sensification system is its *situatedness* (see Section 2.3): Sensifications can either occur where a geospatial phenomenon is actually located, i.e. in their direct real world context—this will be called *location-based*—or in some form of simulacrum that can be displayed on-demand, even if the user is nowhere close to the real location. Here, we have previously distinguished between *scanned* environments and *simulated* environments.

How can we represent this in the diagram? Because there are other outside factors (as we will see later), we can wrap multiple factors into this concept, by finally establishing a concept that already appeared in Section 2.3 and at multiple points in Chapter 3: geographic scenes, as opposed to objects and layers.

4.2.4 Scenes

A sensification is always defined in the context of a certain scene. A sensification system could include multiple scenes, but there needs to be an explicit way to switch between them. With these basic facts established we can ask the question: What is in a scene?

Apart from their situatedness, scenes are defined by the raw data contained within them. Just like in GIS applications, where we have a base map and then the data we want to visualize, we will have a base scene in which the interactive data is placed. We can use Section 3.2 to think about how data acts within a scene and in relation to the sensifications.

The scene itself will always be 3D, i.e. the geometry that defines its extent is a volume. Then, there are one or multiple data sets *located* within this volume—if data is not spatial, it is not something that can be *located in* a scene. If a data set is important to a geospatial sensification in some way, but does not have a direct spatial reference, we need to give it one. And because we are not creating a system that deals in detail with GIS operations and other data transformations, we will assume that this has already happened. The data sets in the scenes are pre-processed to be at the level of quality and spatiality that is required for immersive systems.

Every data set will then have one type: TimeSeries, Trajectory, Coverage, CoverageSeries, Continuant or Occurrent. It will also have a dimensionality: Point, Line, Surface or Volume. For the first five data types we can either try to display them directly, or perform a spatial operation that makes for a more appropriate geometry. Occurrents can not be displayed directly, only by the continuants they affect or by their spatial boundaries. Which data representation and then which channel we choose will determine what the data can be used for. While every set of geospatial data has some form of geometry, temporality and values are optional. This means that there can be sensification systems with no sensory mapping of values. Such scenes are much more focused on interactions and the temporal and geometrical reactions to them. An example in this thesis can be found in Section 5.4. Scenes without temporality on the other hand are common. The data simply remains static over time and only changes in response to certain interactions.

Adding the concept of scenes and their data into our diagram is simple: A scene has a situatedness and is a container for data sets, which have a name, a type and a dimensionality. These data sets then connect to the data channel in the sensification. This connection can either be direct or include an operation (see Section 3.2). It thus usually results in one of the output types from Table 3.3.

We also need to consider that sometimes, we want to interact with the scene as a whole instead of a certain set of data. This usually happens when invoking enabling or scene operators as opposed to sensification operators (see Section 3.3). This is the actual difference between the two interactions that are part of Figure 4.6 ("Change Time" and "Look at Raster"). One is an interaction on something that is displayed (the precipitation raster in the sky) and one operates on the scene itself (changing the scene time).

Finally, what happens to operands once we distinguish between scene and data? The *Attributes* operand in the scene context would be much more appropriately referred to as scene *Appearance*—because every element of the scene that is not explicitly data just establishes context. In a GIS conception, such interactions are similar to changing the map projection or the appearance of the base map. Table 4.1 shows the terms we will use from here on to refer to scene and data operands respectively.

See Figure 4.8 as a demonstration of what the visual syntax for this looks like. Figure 4.9 shows this applied to the weather sensification example. Note that the swimlane boxes are not usually part of the example diagrams, however in the case of scenes they are used as a proper part of the diagram syntax, to denote which data sets a scene encompasses.

Operand (previously)	Data Operand	Scene Operand
Space	Geometry	Space
Attributes	Attributes	Appearance
Time	Temporality	Time

TABLE 4.1: The concept of the interaction operand from Section 3.3 is split into the concepts of data operands and scene operands, with different terms for each.

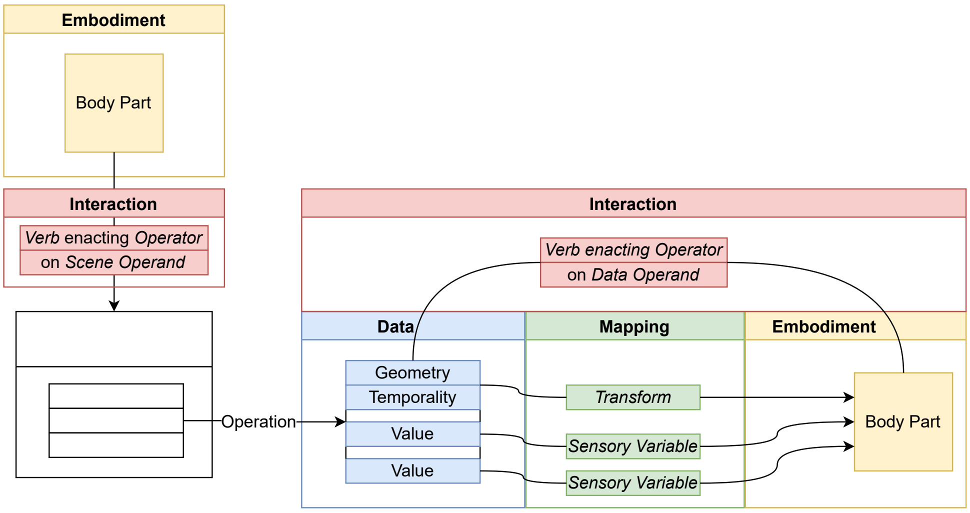


FIGURE 4.8: An alteration of the diagram shown in Figure 4.7. This includes the concept of a scene that a sensification happens inside of, and that scene interactions can be applied to. Data operands and scene operands are distinguished for clarity.

4.2.5 Embodiments and Tools

In Section 3.3 tools were mentioned as an alternative to direct interaction through body parts. This is based on the *extended physiological proprioception* concept, the capability of our bodies to treat tools as extensions of our limbs through their haptic feedback. Tools can thus be seen as an extension of embodiment, i.e. as devices that both can recognize incoming sensory signals, as well as enact verbs.

To increase the clarity of this conception, we can go back to Section 2.2.3, where we established a list of possible interaction devices: Body, Held, Worn, Static, Moving, Moveable. This list includes everything from our own limbs to tangible objects in the environment. If we consider this list, especially objects that are placed in the environments, it is important to realize that movements still originate in our body, and we still perceive their impulses with our body, but the interaction devices act as a sort of interface that allows us to perceive and interact with the virtual environment in different, novel ways.

This means that held, worn, static, moving and moveable objects act like body parts in every way, but they always need to link back to actual body parts. It should be noted that this describes tools that offer new affordances specifically, not just any device—every digital system will use devices to display data or track bodily movements. Most embodiments imply some sort of display system, which should only be mentioned in the diagram if the specifics are important. Some sort of VR HMD for

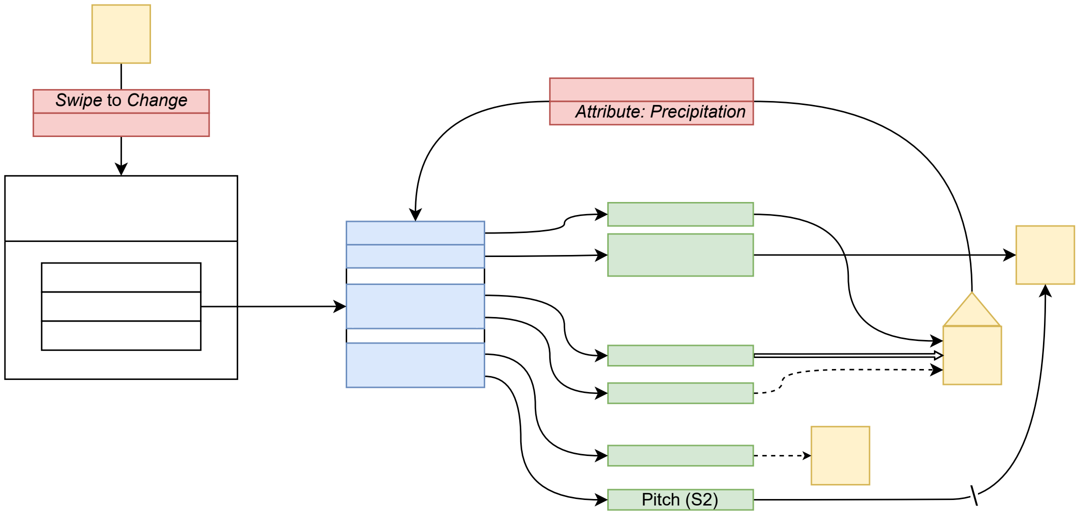


FIGURE 4.9: The example diagram from Figure 4.6, changed in accordance with new concepts introduced in Sections 4.2.3 and 4.2.4. The hand swipe interaction demonstrates scene interactions. The interaction in which a user can "Gaze at" the sky to see numerical values of text is now included in an interaction box with the "Sample Point" operator and the "Attribute" operand aimed at the precipitation attribute.

example is usually implied when a simulated scene needs to be shown to the eyes. A haptic belt on the other hand specifies the actual display location of haptic variables and could be important enough to include in the diagram. What always needs to be included are those tools that enable entirely new kinds of interaction that the body itself can not offer.

Tools can also be entirely virtual simulated objects. An example of such a tool is a virtual baton that can be used to interrogate a sonification in the same way a conductor conducts an orchestra. This virtual tool could even be made to feel like a real baton by a force-feedback enabled haptic glove with shifting weights.

The last possibility offered by tools is to combine multiple body parts. A lens that can be held in front of the eye can enable visual as well as haptic feedback to the hand, or any number of other close-range sensory representations. Showing all these aspects of connections between tools and body parts in the diagram is relatively simple, and shown in Figure 4.10.

4.2.6 Encoding Purpose

The two aspects of user interaction as discussed in Section 3.3 that have not been included yet, are the interaction enhancements and the interaction goals. Enhancements determine what sort of cognitive process over the data the individual interaction is trying to aid with. Goals are part of the whole sensification and can thus determine the purpose of multiple interactions as well as sensory mappings. A sensification either has the goal of perceiving, procuring, predicting or prescribing.

How do these apply to the diagram as it stands? A first interesting observation is that the enhancements have some points of intersection with the value types and sensory variables, even though they stem from the interactivity section. An interaction that aims at a *rank* enhancement for example would need to affect some sort of

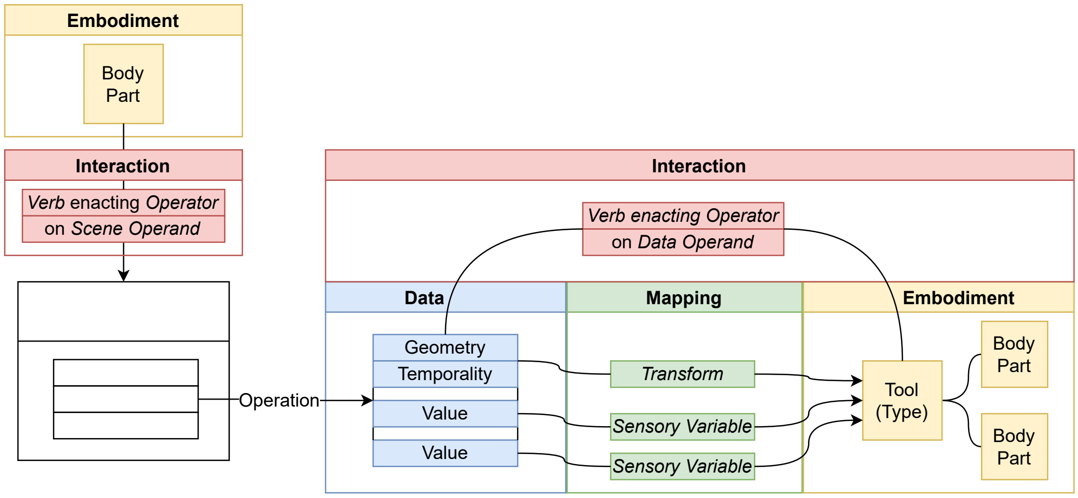


FIGURE 4.10: A version of the diagram from Figure 4.8 that includes the possibility for embodied tools. Tools act like body parts, but they have one of the following types: Held, Worn, Static, Moving, Moveable. Multiple body parts can interface with one tool, and multiple tools can interface with one body part.

ordinal or numerical attribute and then use a sensory mapping that is appropriate for at least ordinal values. The enhancements in the context of the diagram thus are more about what happens in response to an interaction, than what the actual interaction consists of. This makes the enhancement the underlying narrative behind a full sensification feedback loop instead of just an interaction. It is tied to one specific loop in our diagram, originating from the body, traveling through the data, and back to the body. This means that even passive sensification signals that are always-on could have an objective in the sensification—in a way, the interactivity in this case is simply that the user has gone to the trouble of activating the sensification system.

The goals do not intersect with any parts of the other taxonomies in the same way that enhancements do—a goal only describes the motivation behind the enhancements. A sensification that needs to **procure** information will at least need an interaction that *examines*, while a sensification that wants to **prescribe** most likely needs some interaction that *predicts* (see Table 3.5). This is however far less deterministic than other aspects of this taxonomy—a design exercise based on experience more than something that can follow clear guidelines. Thus, the goals are best seen as an outside factor that determines what enhancements need to be in the diagram. They would thus be part of the scene instead of the sensification loop.

The nature of enhancements being tied to certain feedback loops however puts into question the arrangement as it is seen in the figures of this section. None of the figures contain a diagram with separated feedback loops—they always attempt to encode a full sensification system into one loop. To illustrate: the weather sensification example tries to *predict* local weather conditions from an egocentric perspective. However, it contains feedback loops over both precipitation data as well as wind force, which despite the fact that they are different processes that try to enhance environmental awareness in different ways, are not separated. Using multiple smaller diagrams, each with individual goals and enhancements could be the better way to go about this—a concept that will be part of Section 4.3.

4.2.7 Sensing Time and Space

So far, the only step in the diagram that references the geometry and temporality of a data set, is the possibility to apply spatiotemporal transforms, i.e. to change the output type of the data before it arrives at the senses, such that it can be displayed in the ideal way. However, there is more to these aspects. First, there are clear constraints in what kinds of geometries and temporalities each sense can perceive. This does not only act as a constraint in which elements can be connected in the diagram, it also has predicitive power about how applications have to function. If we for example want to sonify a surface raster data set, we need a way to turn the surface into points, because hearing can only accurately resolve point sources. This could work by either sampling the raster spatially (i.e. do a spatial transform), or by sampling it temporally (a process called *temporalization*), usually through an interaction. Vision can perceive higher geometries like lines and surfaces, but can generally not perceive the inside of volumetric objects. These aspects are shown for each of the senses and for both geometry and temporality in Table 4.2.

Sensification	Geometry	Temporality
Visualization	Point, Line, Surface	Structure, Movement, Geometry
Sonification	Point	Structure, Movement
Haptification	Point, Line, Surface, Volume	Structure, Movement, Geometry
Olfaction	Point	None
Gustation	None	None

TABLE 4.2: This table shows which sensory modality of the human body can resolve which kinds of geometry and temporality in real or virtual objects.

Perhaps controversial is the capability of haptification to perceive volumes and internal value structures. This is of course impossible for many common objects that are completely rigid, but data objects do not have to be. Data objects can be moldable, pliabe, liquid, gaseous—this changes nothing about how they interact with the other four senses, but it absolutely allows us to interrogate their internal structure by touch.

We can also circumvent some limitations, like being able to see volumes by employing cross-section views or smart use of opacity. If we use such a technique to circumvent limitations, it is often one of the most important design decisions in the sensification and thus needs to be included in the diagram as either a spatiotemporal transform or sensory mapping.

For the spatiotemporal transforms, we can devise two cases for mappings here: First, the true mapping, in which the sense is naturally capable of perceiving the spatiotemporality of the data. An example is shown in Figure 4.11.

Figure 4.12 shows a transformative mapping, in which the data is transformed with a spatiotemporal operation such that a sense can perceive it, either because the sense originally could not, or some aspect of the system makes the true mapping impossible or too unreliable.

It can be necessary to still include a transformation in a true mapping, because the abstract geometry and temporality does not cover every single aspect a reader needs to know about the representation. There are in fact at least two more factors we have to watch out for: the *range* from which we are trying to perceive a data set, and the *scale* at which we try to perceive it. Figure 4.11 contains an example

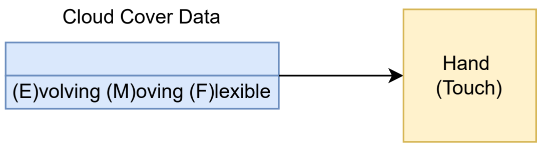


FIGURE 4.11: A single connection from a data output to a body part, as it might appear in the diagram syntax shown in Figure 4.10. This example represents a *true mapping*, i.e. a mapping in which no transformation step (green box) is necessary.

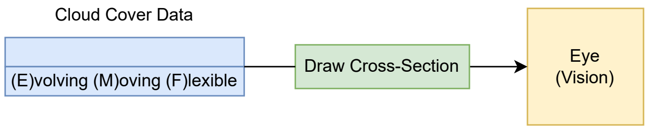


FIGURE 4.12: A single connection from a data output to a body part, as it might appear in the diagram syntax shown in Figure 4.10. This example represents a *transformative mapping*, i.e. a mapping that changes the output data in time, space, or both, in order to enable a sensory organ to see important aspects of the data.

for which this is the case: we might be able to touch an evolving, moving, flexible volume with our hands, but not if it is at the distance or the scale that clouds are normally at. The way we perceive scenes is defined by our sensifications and by how the geospatial objects are related to us in *scale*, *range* and *situatedness*.

4.2.7.1 Scale

Scale can be a very complicated aspect of cartography and remains ever important in immersive environments (see Section 2.3). There are many delicacies to how to scale a virtual environment and the data within it in relation to the user. However, at its core there are only two directions to this aspect: up or down.

We can be at a true 1:1 scale in relation to the geometry of the data, we can miniaturize that data or we can miniaturize ourselves. If the situatedness of a scene is location-based, our scale will always be the 1:1 scale, as the real world is there as a reference for any virtual data that will be displayed. Only scanned or simulated environments can be scaled up or down. And because we are working with geospatial data, all of our data always has the same reference, meaning that all data in a scene will always have the same scale. There can be rare exceptions, like a fully virtual model of the environment that gets partly superimposed on the real world and can then be moved around during interactions, or a grasping motion that pulls full-scale data features closer and makes them smaller, changing their scale temporarily. Cases such as this should sometimes be treated as separate scenes within a scene (as is the case with the full virtual model) and sometimes as the temporary effect of an interaction.

Considering all this, the scale of a scene is thus a given long before the sensification loop starts up, but can be changed for specific sensification loops or interactions. It thus needs to be part of the scene definition in the diagram in the same way the data sets and situatedness are (see the white box in Figure 4.10). This would be

Sensification	Expected Range	Main Limitation
Visualization	Far	Obstacles
Sonification	Short to Far	Distance falloff through medium
Haptification	Touch	None
Haptification (Thermal)	Short	Distance falloff through medium
Olfaction	Short to Far	Dispersion in air
Gustation	Touch	Requires insertion into mouth

TABLE 4.3: A table showing distances at which senses can be expected to recognize signals that originate from a real-world source. This can inform user expectations in an immersive sensification. The three distance categories are *Touch*, *Short* and *Far*.

defined in relation to the user: *True Scale, Miniature Scale, Oversized Scale*. Any interaction that changes scale does so on the scene and in relation to the original scene scale.

4.2.7.2 Range

Intimately important to each sensification loop is its range—giving the user powers akin to telekinesis can be a powerful embodiment in sensifications. Every sense has a specific range that it can register objects at. The range of *visualization* is practically infinite, given enough time and signal energy. *Sonification* is dependent on the intensity of a sound at its source. A signal is attenuated in multiple ways as it travels through the air and other objects and the falloff is much sharper than with light waves. A categorization is shown in Table 4.3. In virtual scenes these limitations can of course be overridden, but the expectations of the user should still be taken into account.

But not only senses have a range, interactions do too. While in the real world, any interaction not involving tools is a touch-interaction, virtual environments permit us to operate on the environment with our body at arbitrary distances. The range of interaction can be different from the range of sensing, even in the same sensification loop. We could perceive a number of objects from a long range visually, but then be required to move closer to them to initiate a touch interaction. Range must thus be either implicitly or explicitly part of each spatial transform and of each interaction.

For the spatiotemporal transform, range can also include how the data attributes are supposed to arrive at the user. This can be a very artificial "moving closer" of the geometry of feature of interest, but it could also include more natural aspects like the "wafting" of smells and the diffraction of sounds. Sometimes this is just a fact of the world-simulation necessary to build a virtual environment, but such signal movements could also be strengthened or break physical reality completely in order to better sensify facts about far-off data.

Considering all these changes, Figure 4.13 includes a diagram that features a similar example to the earlier weather sensification examples. Here however, we track individual clouds as continuants instead of precipitation as a surface, and we want to be able to pull these clouds closer and then visually examine their values. For this, we need a far range interaction on the cloud geometry, and then a way to display the volumetric data of the cloud visually. The diagram also includes the goal of the sensification in the scene, as well as the enhancement of the sensification loop in a comment-element. The values and their mapping are left abstracted.

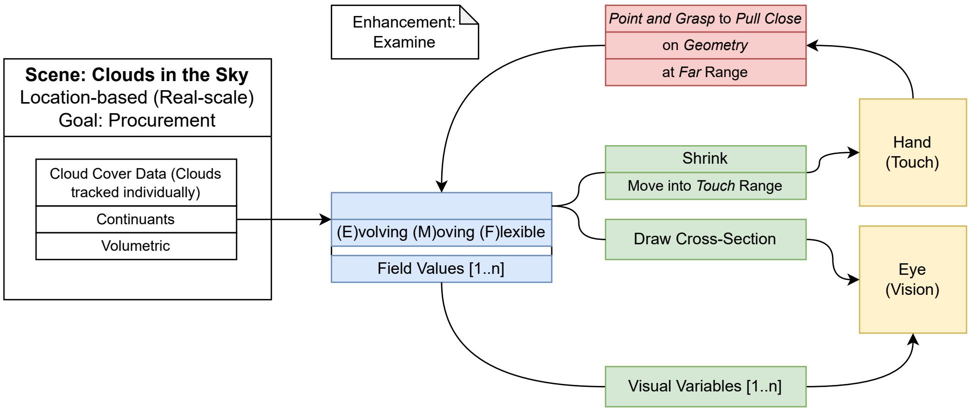


FIGURE 4.13: An example diagram for a system that allows a user to grasp and examine clouds from the sky. Includes elements as defined in Figure 4.10, as well as scene scale, goal, enhancement, two transformative spatial transforms, and an interaction with a range element. Message types excluded. Attribute values and sensory mappings are kept abstracted. The [1..n] denotes that as many attribute as needed can be sensified into as many visual variables as required.

4.2.8 Conclusion: Connecting the Taxonomies

Before we move on to the last step of building our model and diagram language, let us look back on what was achieved in this section: The sensory variables from Section 3.1 were worked into the diagram as a separate class of attribute values whose type and/or purpose can be defined as numeric, ordinal, nominal, selective or associate. Temporality was worked into the same diagram as one aspect of spatiotemporality (geometry being the other). Both temporality and geometry have a clear scope of possible values according to Section 3.2. The way mappings, data and interactions on them were conceptualized was brought into line with how they are actually used in real applications, based on Section 3.3. This is the main point at which all the Sections of Chapter 3 connect. From there, the conception of Embodiment was extended for tool use. A clearer distinction of the scope of interactions in the diagram was established. Then, the purpose of each part of the model was clarified according to the enhancements and goals from Section 3.3. Finally, the relations between senses and the spatiality of data was extended.

What remains now are solutions for scaling of the diagram to larger systems, an improved conception of sensory message types and a clear path to using the diagram for the design of practical embodied geosensifications. All three of these issues will be solved by applying the superpower metaphor from Willett et al. (2021), as first introduced in Section 3.3.

4.3 The Geo-Sensification Diagram

At the beginning of this thesis, we invoked the UML diagram language and how useful it can be for defining systems for technical implementation. What does UML do to be so useful? It separates different aspects of a system into individual diagrams and then gives interface points (for example through package diagrams) to allow us to conceptualize how the different behaviours, data structures, and actors in the system work together. Is something similar possible and useful for sensification systems?

Willett et al. (2021) postulate that we can find inspiration for how to build immersive visualizations in how superpowers are conceptualized in media and popular culture. They deem the concept of thinking about visualization in this way "empowerment", in that the main aim is for users to gain the power to see the invisible and to augment their cognitive abilities, i.e. to empower their perception and thinking. This approach does not start at perceptual experiments or implementation frameworks, as the field of spatial computing often does, but at a design fiction and futuring approach.

This does not come out of nowhere, as the capability of spatial computing technology to give users abilities that feel like superpowers has been explicitly stated by several prominent technologists in the field. And while certain types of powers would require technology that is yet far beyond us, the powers that Willett et al. (2021) focus on seem much more achievable: They call them *epistemic superpowers*, i.e. powers that allow those who have them to "gain knowledge about things, people or phenomena". These powers are opposed to *pragmatic superpowers* which are powers that "actively manipulate things, people, or phenomena" (physical) or "influence thoughts, ideas, and emotions" (mental). As already discussed in Section 3.3, the authors exclusively focus on the visual side of these powers—either what the characters see with vision-based powers, or what the reader/viewer sees when the power is used.

What is different in our case, is that we build the world in which we want to have these superpowers. Instead of learning information *about* existing things, people and phenomena, we can turn our data into these things, people and phenomena, and learn about our data *through* our interactions with them. This means that at least in scanned and simulated scenes (i.e. most VR applications) we can also utilize physical pragmatic superpowers as tools of epistemic analysis. Considering this, what conceptual tools are contained in this notion of *empowerment*, and can we use them here?

The authors define seven axes of empowerment that a superpower can induce in users. These axes are scope, access, spatial relevance, temporal relevance, information richness, degree of control, and environmental reality. They act as measures of quality, similar to the concepts of expressiveness, effectiveness, and appropriateness in visualization (see Section 2.1). Because some of these might either be a given or not relevant for geosensification in particular, we will consider their use as a quality measure one by one:

Scope is extremely reliant on the application context in which we think of powers—geospatial sensifications can be a lot more specific than the average superpower and still be useful in the context of geospatial analysis. As such, scope is a less important measure of empowerment.

Access is a very important design consideration—the more specialized hardware a geosensification needs, the more it will be a specialist's tool. In the worst case, even an extremely useful geosensification will never leave the lab purely because of accessibility issues.

Spatial relevance will always be high in geographic scenes, as all geospatial data is exactly related to a location.

Temporal relevance is important both for data that has temporal dynamics, as well as in static data sets—quick feedback to interactions and timely sensory representation is very important for a power to feel engaging, however it is best renamed to *responsiveness* in order to not confuse it with the temporality of the data.

Information richness refers to factors like the "quantity, variety, and accuracy of the information it is able to convey". This is perhaps the most relevant axis of empowerment for geosensifications.

Degree of control is very important, as without rich interaction within the geographic scene, a geosensification is just a multi-sensory map.

Environmental reality places, perhaps controversially, situated visualization as inherently more empowering than fully virtual environments.

That situated scenes are broadly more empowering than scanned or simulated scenes might well be true if we consider the use of an epistemic superpower in the real world in comparison to using it in a virtual environment. However, fully virtual environments of course allow us to switch between multiple environments at impossible speeds or even turn the data itself into wholly new environments. Still, it is an important factor that stands at the start of each geosensification, especially when considering how additional sensification objects like tangible objects can augment an environment.

To summarize how relevant these axes are for our context: spatial relevance is a given and scope is determined by the exact application. This leaves access, responsiveness, information richness, degree of control and environmental reality as the five main design pillars that allow us to enhance the feeling of empowerment in a geosensification.

There are many more considerations in the original paper. However, more than all the details laid out here, the most important conceptual tool in the article is the simple idea of seeing visualizations as one or multiple distinct superpowers. This ties in perfectly with the sensory feedback loops as presented in this thesis so far. Such a feedback loop always involves one or more interactions and one or more spatiotemporal or sensory mappings, but it is a distinct loop in regard to the rest of the system. This allows us to separate multiple loops into individual "powers". Each of these powers then have a number of enhancements they try to give the user. In conceptualizing this, we could either start from the bottom up, starting with an idea for a power and see what kinds of sensification enhancements they could offer us, or we could start from the top down with a list of goals and see what kinds of enhancements need to be offered so a user can attain their goals.

This conceptual tool of sensification loops as superpowers also gives a stronger basis for last part of the original diagram draft in Figure 4.2 that has not yet been updated: the four message types of queried, feedback, interrupt and continuous. When considering sensification loops not as messages containing data values that are sent to the user, but as superpowers that are empowering the user, this can be verbalized far more easily: as the activation condition and duration of a superpower.

The possibilities here as adapted from the message types of Berger (2021b) are *instant* (activated once when needed), *sustained* (activated and kept on during interaction), *reactive/conditional* (activating themselves based on some condition or in reaction to some event) or *constant* (unable to be switched off, for good or ill). We can use this in the diagrams as-is, by defining this as the "Activation" of a power. The activation is defined on a per-power level. If a power is reactive/conditional, we also need to define the condition for it on the per-power level.

This introduces one more problem: A power might involve multiple sensory mappings that react differently to possible interactions. It could therefore be important to define for a sensory mapping which interaction it belongs to. This can be included in the diagram by naming the verb of the embodied operator below the sensory variable.

With this, the integration of the taxonomies into a model and diagrammatic representation is complete. The final version of the diagram is shown in Figure 4.14, for a slightly altered weather sensification example similar to the one shown in Figure 4.9. The three aspects that used to cause ambiguities between the spatial transforms and sensory mappings are now clearly separated into three different powers. These powers have been given names similar to what superpowers in media and popular culture might be called. Tracking these metaphors throughout the design process can help both the designer in maintaining a clear conceptual reference for a power, and it can give first time viewers a chance to quickly recognize the basic idea behind the system.

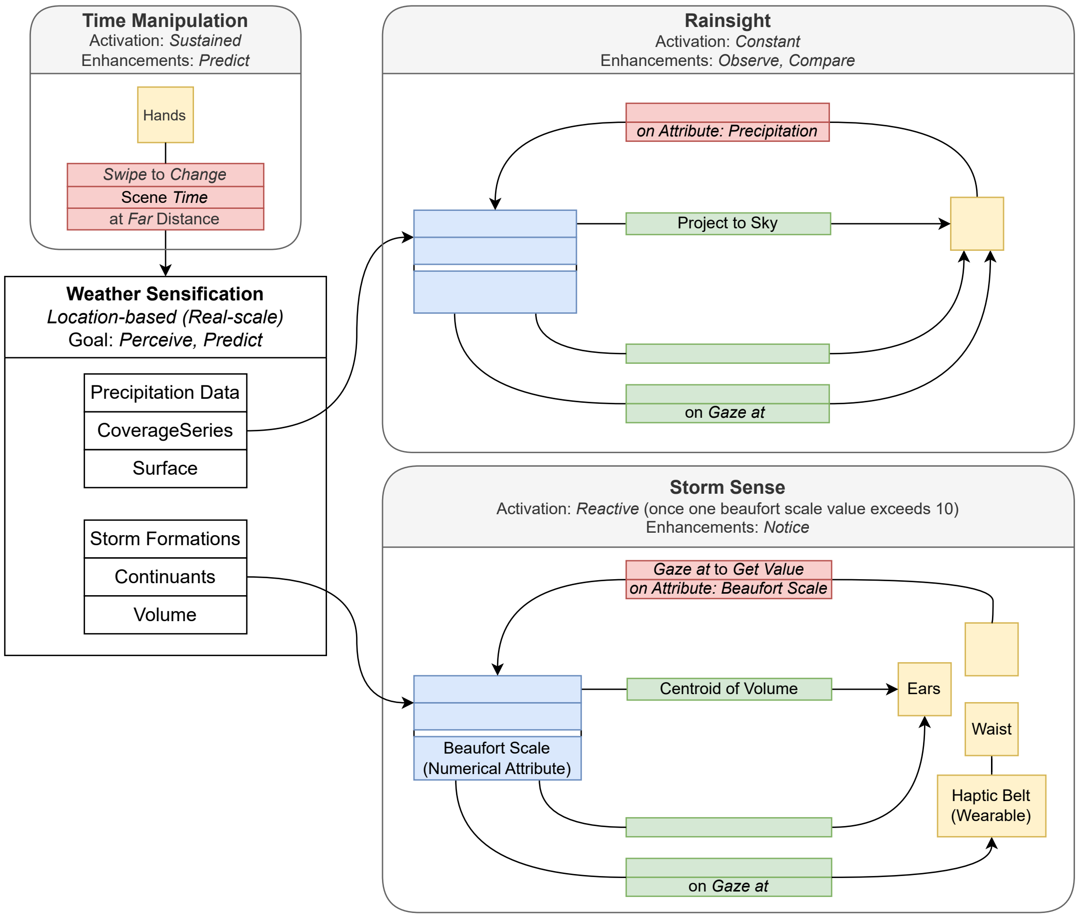


FIGURE 4.14: A rework of Figure 4.9 into the new power-based conception of embodied geosensifications. The different message types were removed in favor of power activation types. Goals, enhancements, scene scale and interaction range are properly defined. Note that we now consider the beaufort scale an attribute of a storm continuant instead of a field that is part of the weather surface, to highlight how multiple powers with different data sources operate towards a common goal in a scene.

4.4 The Steps of a Sensification

Because this model has many moving parts and possible interactions, it can be helpful to have clear list of questions to answer and decisions to make when designing an embodied geosensification. The following list is an attempt at such a tool. By going through it and considering the points step-by-step, a complete diagram in the same syntax as introduced in this chapter can be constructed. Note that this is not the only possible order in which to ask these questions.

1. Establish one or more scenes. (Section 4.2.4)

   (a) Consider what situatedness is required for the scenes: Location-based, scanned environment, simulated world

   (b) If the scene is a scanned environment or simulated world, consider its scale: True Scale, Miniature Scale, Oversized Scale

2. Define the required data sets for building each scene (see Section 3.2). For each data set ask yourself:
   (a) What is its dimensionality in the scene: Point, Line, Surface, Volume?
   (b) What is its data type: TimeSeries, Trajectory, Coverage, CoverageSeries, Continuant, Occurrent?
3. For each scene:
   (a) Consider what the goals of the scene are, i.e. what kind of knowledge do you want to gain from the data contained in it: Perceive, Procure, Predict, Prescribe (Section 3.3.2)
   (b) Consult Table 3.5 and consider what cognitive processes (enhancements, see Section 3.3.2) are required to fulfill these goals: Observe, Notice, Examine, Compare, Rank, Associate, Enumerate, Measure, Predict, Delineate
   (c) Take the enhancements and split them into one or more powers. For each power:
      i. Consider possible metaphors for how these powers operate.
      ii. Select one or more subsets of data from the scene. For each of these subsets, consider how it needs to act in time and space (Section 3.2) in order for the enhancement to work, and whether an operation needs to be applied to it to prepare it for data display:
         A. What is the geometry that needs to be perceived: Point, Line, Surface, Volume?
         B. Does the geometry move over time? (Stationary or Moving?)
         C. If it is larger than a point, does its internal structure change? (Uniform or Evolving?)
         D. If it is larger than a point, does it change shape? (Rigid or Flexible?)
      iii. Pick senses that you want to display the geometry and temporality of your data with.
         A. Consult Table 4.2 to see whether this sense can directly display your data geometry and temporality.
         B. If not, define necessary transformations in geometry and time (see Section 4.2.7).
         C. Consider the spatial relation between user and data and whether there need to be range and scale transformations to make the data accessible (see Section 4.2.7).
      iv. Pick the values from your data set that are required for the sensification to display that it needs to display. For each attribute:
         A. Figure out as what type it needs to operate (see Section 3.1): Numerical, Ordinal, Nominal, Associative, Selective
         B. Figure out whether the value acts as an attribute or a field over the geometry.
         C. Consult Table 3.1 to see which sensory mappings work for the attribute values depending on your objectives and how technologically feasible they are. Temporal and spatial variables can only be used if the temporality and spatiality of the object is static in some way.

D. Consider whether you already have an appropriate geometric reference to display values to the user (see Section 4.2.1). This can be a reference of a different sense.

v. Consider how and when this power should be activated (see Section 4.3): Instant, sustained, conditional, constant

vi. If conditional, specify the condition.

vii. Consider whether the required enhancements are already fulfilled by just the sensory mappings on their own.

viii. If not, or if your activation type is instant or sustained, introduce interactions (see Section 3.3). For each interaction:

A. Consider whether the interaction needs to target the entire virtual scene, or just one data set. Choose an operand (see Section 4.2.4). If you are interacting with the scene, choose one of: space, time, appearance. If you are interacting with a data set, choose one of: geometry, temporality, attributes.

B. Pick an operator that allows you to reach the enhancement.

C. Choose a body part and which movement (verb) this body part needs to perform to activate the operator.

D. Consider whether there is a real or virtual tool that would fit the interaction more than a body part (see Section 4.2.5). Consider which body parts of the user interface with the tool. Pick the type of tool: Held, Worn, Static, Moving, Moveable.

E. Consider what range the interactions happens at (see Section 4.2.7) and if the body part can naturally reach the features. If not, add a mention about what range the interaction is supposed to operate at (i.e. if the user is supposed to walk towards the feature, or have something akin to "telekinesis").

F. If the interaction was included based on a instant or sustained power, note the verb down with the appropriate sensory mappings (see Section 4.3).

ix. Check if the powers fulfill their enhancements, and if the enhancements fulfill all scene goals.

4. Consider the quality of your sensification idea (see Sections 4.3 and 2.1):

(a) Expressive: Does this sensification display data that is actually relevant?

(b) Effectiveness: Can we actually perceive the sensory mappings well?

(c) Appropriateness: Is the difficulty in creating this sensification within reason in comparison to the usefulness it can have?

(d) Accessibility: Is the required hardware reasonable for the target audience?

(e) Responsiveness: Is the sensory feedback loop fast and reactive enough to be engaging?

(f) Information richness: Does the chosen sensification display the full quality and richness of the data set?

(g) Degree of control: Do the interactions enable the user to deploy powerful forms of analysis?

(h) Environmental reality: Does the sensification use the given environment it is supposed to work in to its fullest extent?

Figure 4.15 then shows a graphical overview of this process, highlighting how decisions about different aspects of the diagram tend to influence and follow from each other.

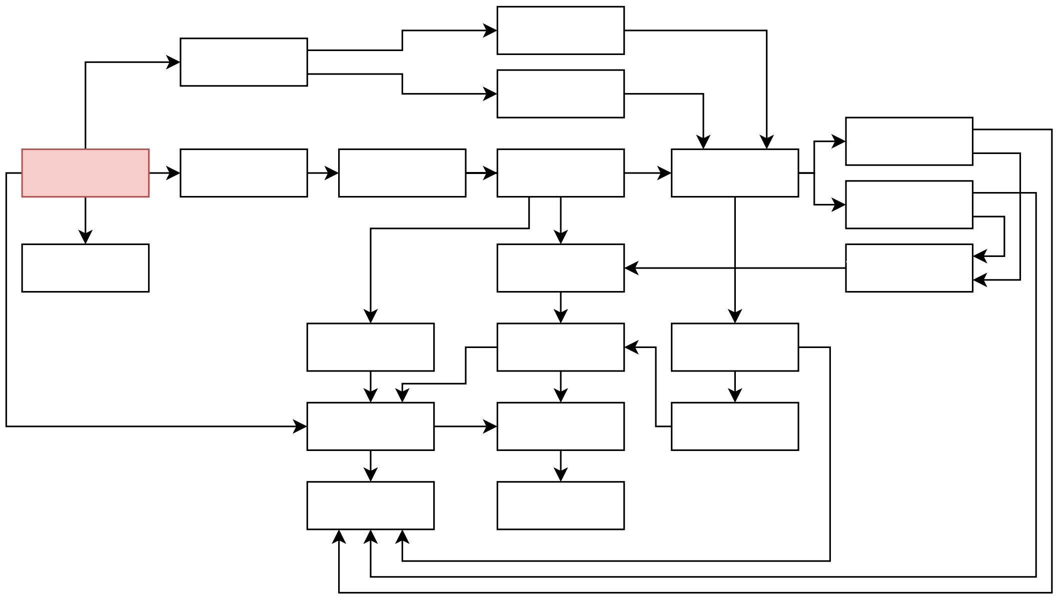


FIGURE 4.15: A high level overview of the different design decisions in creating a embodied geosensification and how they depend on each other. Following the guideline in Section 4.4, we start by defining a collection of scenes (element marked in red). Decisions about range and scale of scenes, interactions, and data, as well as quality considerations removed for readability.

Chapter 5

Case Studies

5.1 Introduction

The motivation to establish these theoretical guidelines is rooted in several years of working on practical problems in immersive analytics applications. During these years, multiple working prototypes have been developed. Some of them reached a maturity fit for publication. The geospatial data underpinning each prototype is quite different, as the possible application areas of geosensification are quite varied. There are some conceptual through-lines however—the main one being that they all in some form deal with data sets who have surface geometries.

Every single type of geometry has its issues when employed in immersive environments, but as we will see, surface data sets illustrate the core problems quite well. First, there is Berger and Bill (2019), which laid the groundwork for this thesis by involving multiple senses in a simulates scene aimed at data representation, specifically in the context of urban traffic noise. One of the problems uncovered during work on the prototype was the difficulty in visualizing surface data sets for the sort egocentric perspective common in immersive systems: Once we are standing *on* a data surface, the readability quickly degrades towards the horizon. In the prototype this was successfully solved by allowing the user to change scale. But as established in Chapter 4, this becomes impossible in situated analytics, and has large implications for what kinds of interactions can be employed, because we increase the range to the data set. Another problem is how the spatial resolution of GIS data interacts with an immersive display context: a continuous raster data set suddenly appears like a vector data set of very sparse discrete points if we are at a true 1:1 scale.

Berger (2020) tried to solve the first problem by moving the display of the surface data fully to sonification. It became clear that only because a sense has a capability (surround, long-range hearing) it does not mean that it is trivial to make a sensification that plays to this strength. A simple pitch-based mapping failed to be sufficient for sonifying the data in the given situation, which led to questions about what other sensory variables were available and how useful they might be. An important reference point in the paper was Dubus and Bresin (2013), which features a complex taxonomy of sonification variables. This taxonomy was in many ways what inspired the research in Section 3.1.

At the same time there was also work on immersive weather sensification, as shown throughout Chapter 4. Here, the main focus was mostly on visualization, with simple attribute mappings. However, the question of geometric resolution remained: When trying to visualize contextual information about formations of rain clouds for situated analytics, how do we even get to the needed quality of data? What is trivial for a birds-eye view (as known from any modern rain tracking weather app) becomes extremely difficult when operating from a human perspective. All available weather data suddenly is extremely coarse both in geometry and in temporality, and it becomes almost impossible to usefully show detailed information. This was the main motivation behind Section 3.2, specifically behind the concept of the continuant. Continuants in many ways show the data quality that is required for immersive sensification system to work, as the data suddenly needs to work as the basis for the highly detailed and dynamic objects we expect to be rendered in real-time 3D applications, instead of for the large-scale GIS rendering most data formats were originally conceived for. This will not be discussed in further detail here, as a finished diagram representation for such a weather sensification was already shown in Figure 4.14.

Application	Senses	Data Types and Scales	Interaction Effector
Sonifying Noise	VS	Line, Surface (Numerical, Ordinal)	Handheld controller
Sonifying Coverages	S	Surface (Numerical)	Handheld controller
Transforming Spaces	VSH	Points, Surface (Numerical, Ordinal)	Hand microgestures

TABLE 5.1: An overview over which senses are sensified to, which data geometries and value scales are involved, and what is the main interaction effector for each use case.

To explore alternative solutions to surface data display, the focus turned to an older technique for cartography with an egocentric perspective: multi-perspective views, as introduced in Lorenz et al. (2008). These views are a working solution to immersive visualizations in which we imagine ourselves standing on a 2D map with height elements (like for example a city model). The map is curved upwards the farther it gets from the viewer, filling the horizon with a map representation that would usually be occluded. However, immersive environments are not always this reducible to a plane. They can be complex, temporal, volumetric, and sometimes labyrinthine. In order to fully explore them, one needs to be able to move through the environment but also control its appearance and configuration—in short, one needs a strong set of interactions. After much work on sonification, first trials with haptic interfaces and the body movements inherent to any situated analytics application, the concept of embodiment was not far. The idea to break out of the very conventional interaction paradigms in previous work resulted in the research that led to Berger (2021a), which explored what it could mean for multi-perspective views to be embodied and deeply interactive.

In this chapter, we will now use these systems to see how they could have been represented through our model and diagram, applying some of the lessons that were learned along the way. Table 5.1 shows a an overview of which aspects are involved in the created powers. They do not use the whole space of sensory or interaction possibilities, but combined with the weather example, all aspects of the diagram are shown at least once. This is also an opportunity to validate that the diagram can describe systems that were not specifically created as examples for it. For the first example, we will go through the full list of considerations established in the last section, to demonstrate the process. Later on in the chapter, we will only consider aspects as they become important.

5.2 Sonifying Noise

5.2.1 Motivation

Noise emissions caused by traffic in urban environments are one of the current issues in the fields of urban planning and public health. Since 2009 the EU has started developing and adopting the "Common noise assessment methods in Europe" (CNOSSOS-EU) noise policy and calculation model (Kephalopoulos et al., 2012), aimed at creating a common standard for noise mapping for european cities. Noise maps show how traffic noise is emitted from roads through the urban environments surrounding them.

These maps usually show a raster of sound pressure levels adjusted for human hearing (dB(A)), averaged over a number of propagation and weather scenarios. This allows an efficient evaluation of problematic areas and how well a city is dealing with urban noise in comparison to others. The issues however are, that noise as it arrives at us has many more complexities than just being sampled as a point on a surface, and that dB(A) follows a logarithmic scale. These maps are thus tools that an expert can read, but it can be difficult for a layperson to understand what exactly the colored raster values mean. Factors like frequency modulation over distance and different types of grounds, diffractions and reflections can also influence what exactly the noise sounds like in a certain space, which is only partially captured by the calculation model. Sometimes it would be more appropriate to actually *listen* to what the noise map implies.

This basic idea was the drive for developing the solution presented in Berger and Bill (2019): To offer an immersive method to explore urban noise maps. Since the original paper was published, such sonification methods have appeared in professional traffic planning software (though not yet in an immersive way).

5.2.2 Technical Summary

The original technical solution starts with a CityGML data set of LoD (Level of Detail) 2, as would be available for many European cities today. This data needs to be brought into a real-time 3D engine like the Unity game engine (as current VR technology is built almost exclusively on a game-focussed technology stack, not on GIS). At this LoD, these data sets contain geometries representing the boundaries of buildings and other urban structures, as well as additional attribute data. While this in concept makes them perfect for such an application, there are still issues of data quality. Not only does the topology of the data need be perfect (i.e. the ways in which individual elements fit together), the geometries need to be put into relation with a sufficiently precise digital elevation model (DEM). In theory this should be solved by the fact that all data has a proper spatial reference, however in practice data is of course often imperfect.

Additionally, there is the issue of transferring this spatial data based on double precision floating-point coordinates into the single precision of the Unity game engine. This transference is simple for the building geometries, which use projected coordinates and can thus simply be moved from the global coordinate system into a local coordinate system, yielding sufficient accuracy for an egocentric, real-scale display perspective. It is much more difficult for the DEM data, as projection distortions play a much bigger role. In Berger and Bill (2019), this was solved mostly through GIS-based data preparation and the Mapbox Unity plugin, which at the time was one of the very few off-the-shelf prototypes for including generalized geospatial data into the engine.

Based on this data, a simple simulated scene could be rendered and standardized VR interactions like teleportation included. The difficult part then was to simulate the noise data. The state of the art at the time was to base immersive real-time sound simulations on noise samples recorded in the field. Larger scale noise mapping on the other hand is based on simulation data, which was the focus here. There are several simulation models with different methodologies. The CNOSSOS-EU noise standard brings with it a geometric simulation model. This model is based on ray-like path propagation and thus a fitting basis for implementation into a game engine. These engines have deep libraries for dealing with ray-casting and line collisions, usually employed for use cases like character movement or firing weapons. Berger and Bill (2019) thus took this geometric model and implemented a version of it into the Unity engine. The model defines four types of propagation path—horizontal diffractions, reflections, vertical diffractions, and paths with a combination of vertical diffractions and reflections (see Figure 5.1). Each path uses a slightly different sound attenuation model. For this, a heuristic was implemented that covers all four types of path, limited by angular resolution (the number of paths that are traced outwards from the listener point) and the maximum number of reflections (only one), in order to keep the model usable in a real-time context. This model is then calculated over a regular raster of listener points, which are colored green to red based on the calculated db(A) at that point. Points distributed over the line geometry of the roads act as sound sources.

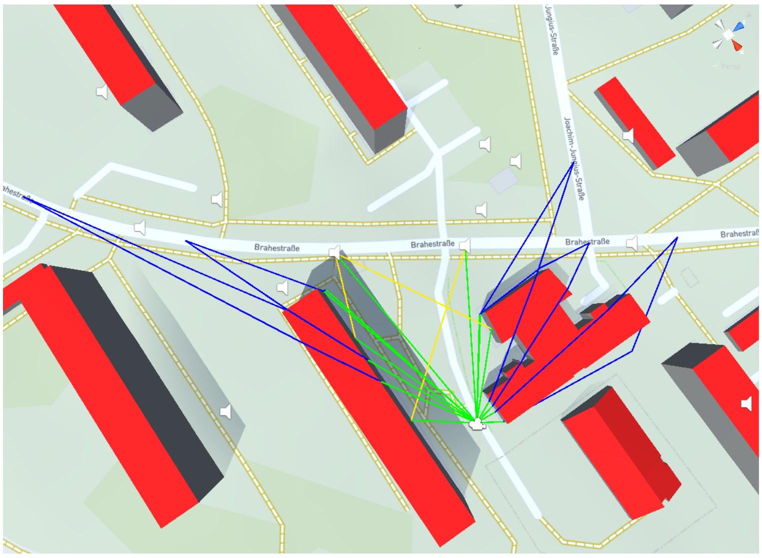


FIGURE 5.1: "Noise propagation in a street with seven source and one observer point. Green lines are direct paths and path end segments. Blue represents paths that lead to diffraction around buildings, and yellow paths that cause reflections on vertical building surfaces. The white speaker symbol shows where the image sources would be located if their height information was removed. (Otherwise some would end up above the camera.) The observer is represented by a camera, the source points are located at the beginning of the line segments.", from Berger and Bill (2019).

In addition to this noise-map-like raster, the user has the option to render the noise models as a sonification. This sonification uses the same underlying model, but instead of calculating an averaged db(A) value over all paths, it keeps the direction, length and sound pressure of each incoming path intact. Then, once the user starts the interaction, the same sound of a driving car is played over each path, with its volume shifted according to the db(A) and its start time delayed according to path length. This creates a sound with a strong initial signal, some reverb and late reflections, entirely simulated by directly applying the policy-driven noise model.

Apart from the obvious performance problems of using a complex geometrical model on a 2018 standalone VR device, there were of course issues with this approach. Interestingly, the main issue was not in the sonification, but the noise raster, visible in Figure 5.2. It is there to give overview and context to what the user is hearing, but it can only be properly viewed from the top-down perspective. Otherwise there are occlusions from buildings, as well as self-occlusions and lowering readability towards the horizon. This would only get worse as the geographic scene increases in details (like for example through the addition of trees). This is what eventually led to the research in Section 5.3. Before we resume there, we will now consider what the model from this thesis would look like as applied to this application.

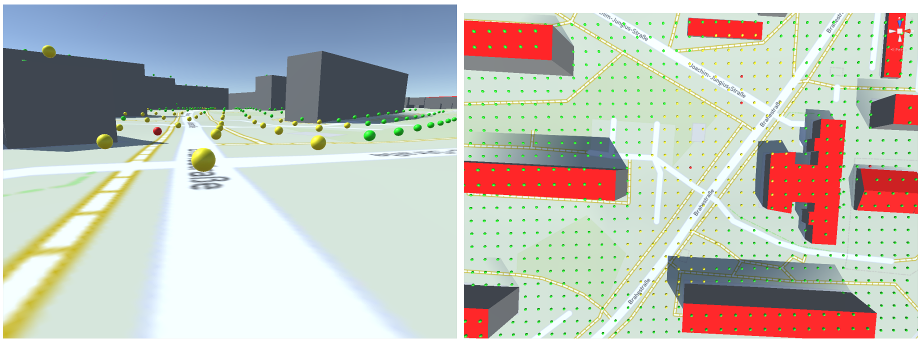


FIGURE 5.2: "(**a**) View on the noise grid from the user's perspective (**b**) Top-down view on the noise grid", from Berger and Bill (2019)

5.2.3 Applying the Model

Following the step-by-step list from Section 4.4, we begin by establishing the *scene*. As-is, the sensification includes a **simulated** world that is at **real** or at **miniature scale**. The required *data sets* are surface data sets for elevation and buildings, as well as a line data set with the sound emission levels over all the roads in the city model. For simplicity we will only include the sound emissions in the diagram, which would be a **line coverage**, because the sound emissions of the road change over its geometry, but the geometry remains static over time.

In this scene, we try to reach the *goal* of **perceiving** the noise information in novel ways. For this, we want to *enhance* the way in which user's can **notice** noise values in the environment, and then how they can **examine** and **observe** them in new ways.

Thus, we split the enhancement goals into two *powers*: In order to notice the values, the user must be able to **see noise**—the metaphor here is simply the concept of a noise map placed on the ground—recognizing db(A) values from far away. Then, we want them to be able to more viscerally understand the noise by being able to **hear the noise**.

For seeing the noise, we can create a data **surface** that displays a **field** of **ordinal values** over the ground—different levels of noise thresholds, i.e. common limits that are set by public health policy. This raster can be **calculated** from the data sets already present in the scene, by applying GIS *operations* between scene and power. In the original paper this surface is treated as **uniform, static and rigid**, however depending on the perspective we average from it could also be evolving instead of uniform. Because we are visually looking at the noise, the noise surface can be displayed directly and no *spatial transform* is necessary. In the original paper the surface is still transformed into points, in order to keep a view on the underlying terrain base map. For ordinal values in visualization, the *sensory mapping* choice of hue (V3) in the original paper is not ideal according to Table 3.1. **Color saturation (V2)** would be more effective.

Because this power is supposed to orient the user and does not otherwise obstruct information (except for self-obstructions), its *activation* can be **constant**. This power now fulfills the **notice** enhancements, i.e. no more *interaction* are necessary.

The second power is the power of hearing the noise. To hear the noise means to hear the propagated noise emitted from the roads. This means that we need to start with a **line** output (that can be **uniform, static, and rigid**) and then apply the model as a *spatial transform* between the line and the user. The model involves a **segmentation** of the road, a calculation of the **propagation** for each segment, and finally a **rendering** of the propagation result as a point (i.e. as a geometry that is actually audible). This also means that we are not sonifying the sound emissions from the road, but the derived **numerical** values of sound pressure (i.e. what arrives at the user position).

Then we have to consider what *sensory mappings* to use for the sound pressure. The simplest solution here would be any variable that is good at representing numerical values. However, we are not actually trying to make the user recognize a number—we are trying to make them recognize whether a noise level is manageable, annoying or unhealthy. This is an ordinal use case, even if the values are numerical. If we limit this to sonification variables, according to Table 3.1 we could use duration (T1), rate of change (T2), order(T3), frequency (T4), composition (T8), location (Sp1), speech (L1), volume (S1), pitch (S2), envelope (S4) or noise (S5). Noise is a misleading variable name here, as the noisiness of a synthesized sound signal is a very different phenomenon compared to urban noise. Speech is not the right choice, as we have many points along the line, and it is not perceptually selective. Because the problem of noise is intimately tied to persistent volume, volume is a better choice, especially because it is also at least partially effective at being selective.

Following this logic, i.e. the logic of a purely abstract mapping of data to sensory variables, rate of change, composition, pitch and envelope would be even better choices because of their higher selectivity. However, an abstract mapping was of course not the original intent—the noise actually needs to sound like urban noise, which requires more than just an abstract noise signal. In order to achieve this, we need to actually **compose** (T8) a sound that viscerally represents urban noise, that sounds need to have the correct **volume** (S1) and **timbre** (S3). Even the influence of the reflections and diffractions through the **environment** (Sp5) and the perceived **location** (Sp1) of origin is a factor of the data sonification here, not just of making the scene immersive. This creates a highly complex sensory mapping, but one that has its roots in physical reality and incidentally features many variables that work well for ordinal, selective display—and should thus immediately be readable to anyone utilizing the system. If the variables of the physicalized mapping were variables with low ratings for ordinal scales and selectivity, this approach to sonifying them would be likely to fail.

Moving further down the step-by-step list, we arrive at *activation*. This power is one that inherently displays an unpleasant fact. As such, it should not be always-on, but played only when a user chooses to. This leaves the instant or the sustained activation. In the implementation as presented in the paper, the activation would be **instant**, i.e. the values are played once when a button is pressed. As such, one *interaction* is required in the power. Because this system was running on the Oculus Go VR headset, the only interaction *tool* was a simple rotationally tracked **handheld controller**. This means that the interaction has to have the user **pressing a button** (*verb*), to **play** (*operator*) the **sound pressure attribute** (*operand*).

Based on this controller, we can also include the two additional powers that were present in the original system: teleportation and scale change. These are pragmatic powers that operate on the scene and we will not discuss them in detail here. All four powers defined according to our diagram syntax are shown in Figure 5.3.

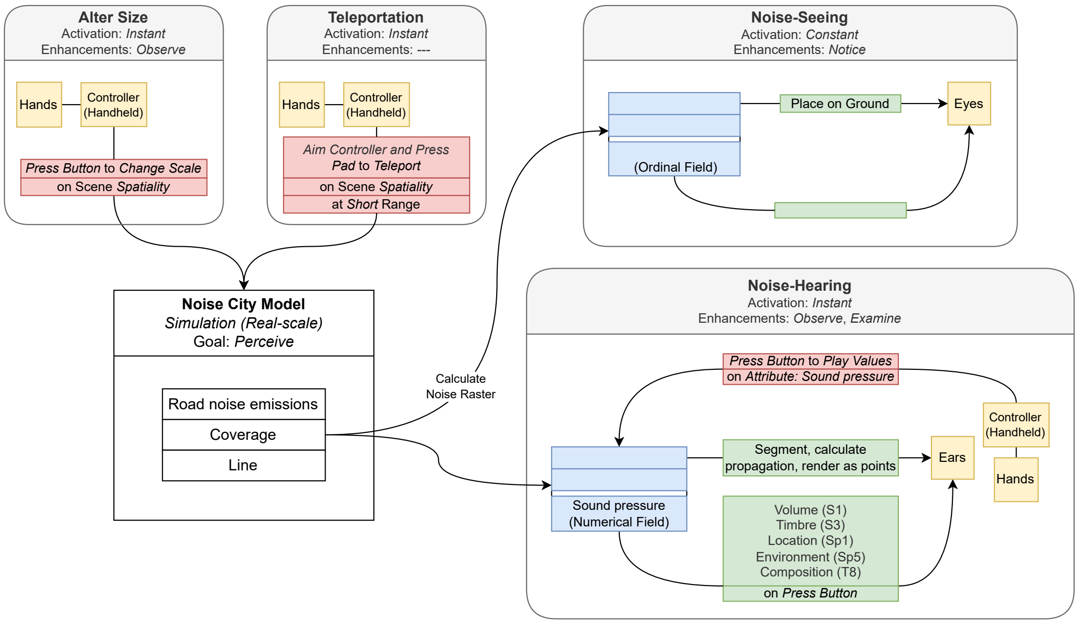


FIGURE 5.3: The diagram syntax from Chapter 4 applied to an updated version of the noise sonification from Berger and Bill (2019). It shows a scene with four powers, two of which contain a sensification loop.

Applying the model in this way of course does not solve the problem of self-occlusion in egocentric coverage visualization discussed at the beginning of this chapter, because the problem of **noticing** occluded values is circumvented through the **alter size** power. However, this power is only possible because we can change the scene scale and the density of scene elements in a **simulated** environment. As soon as we are in a **location-based** context (i.e. one in which we can not change scale) circumventing this issue is much more difficult—which sets the stage for the next section.

5.3 Sonifying Coverages

5.3.1 Motivation

As soon as there is no way to alter our size for a top-down view on the environment, visualizing a coverage reduces drastically in effectiveness. For Berger (2020), we thus strip the noise context away and instead only consider an unspecified **surface** showing a **field** of **numerical** values in a **location-based** scene.

The basic idea of the paper was that by using sonification instead of visualization, these values could simply be made to be occlusion-free, because sonifications are not limited by occlusions in the same way that visualizations are (see Table 4.3). Sonification also operates in a more passive, omnidirectional capacity than vision, making it appropriate for the targeted **notice**-enhancement.

5.3.2 Technical Summary

The basic issue with sonifying a surface, is that sonification can only resolve point values (see Table 4.2). Berger (2020) uses a *temporalization* approach to this issue, i.e. the raster is sampled on a large number of point positions (specifically in a hexagonal arrangement), which are then played sequentially over time, in order to not play hundreds of values at once. Specifically the temporalization uses the metaphor of a circular scan, in which the values are played in concentric rings centered on the user, as illustrated in Figure 5.4.

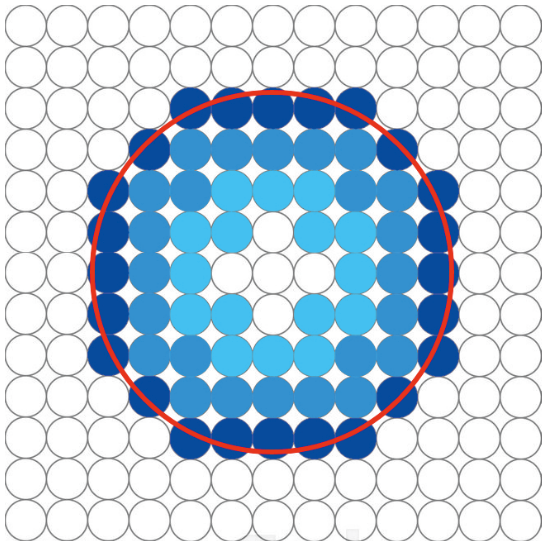


FIGURE 5.4: The principle of temporalization utilized in Berger (2020), in which a regular grid of point positions is played based on a concentric ring (red) that travels outwards from the user.

While this solution works on a technical basis, a user study showed a limited amount of effectiveness. The reason for this was difficult to conceptualize without a clear idea about the perceptual implications of many of the choices that were made during the original implementation.

5.3.3 Applying the Model

Instead of going through the full step-by-step guide, we will consider only individual parts of the model to get a clearer view on the issue. First, in the paper the variable of pitch (V2) was utilized. According to Table 3.1, this is not a very effective choice for numerical display. It does have a good selectivity, which could be helpful for this specific problem, but even after the temporalization, a very large amount of values are played at once. Either a much more effective variable mapping is required, or the values need to be classified and turned into an ordinal scale.

Going by Table 3.1, the most effective sonic mapping for high selectivity and a numerical value would be that of **envelope (S4)**. It is also likely that a high associativity could help, as we want to quickly identify clusters of similar values (the hills and valleys of the data). In order to achieve all three, a multi-variable mapping would be required. If this does not solve the issue, the next aspect that could possibly be improved is the spatial mapping of the data. In order to reduce the number of values drastically, we could focus on those regions that are most interesting to the listener, by not sampling with a regular point grid, but instead turning any local minima and maxima into source points for the sonification.

A different way to approach this would be to change the *temporal* variable—something that the original paper could not make explicit, for lack of a clear conception of temporal variables, was that the temporalization implicitly made use of the variable of order (T3) in how the values are displayed based on their distance to the user. Order however does not yield good associativity or selectivity. A better variable could be **synchronization (T6)**. In context this would mean that all similar values would be played at similar times, which could be implemented as different "heights" in the coverage getting played at different times. This could even enable the sonification to run in the background—periodically, the coverage could be shown by playing all points in the deepest valleys, slowly moving up to the peaks of the data surface. Note that synchronization does not represent numerical values well—this would still be the purpose of pitch or envelope.

The original concept for this sonification also includes the possibility to include it as an optional long-range display in a system that primarily operates at close range. As such, Figure 5.5 shows the synchronization-based option from above and simultaneously implies a larger context by including an unspecified examination-focused power.

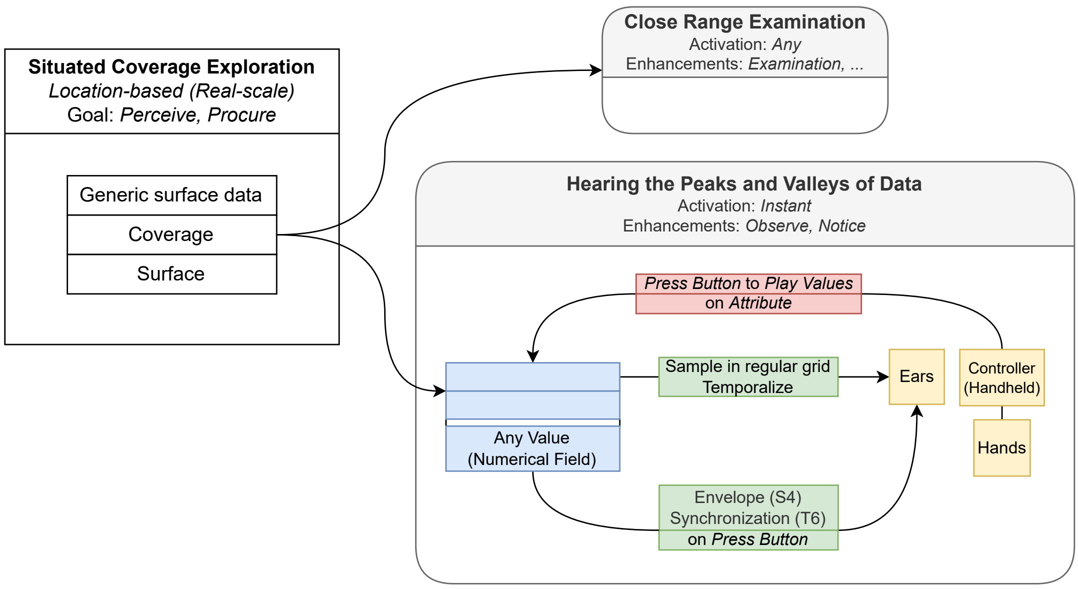


FIGURE 5.5: The model from Chapter 4 applied to a slightly adapted version of the coverage sonification from Berger (2020). Sensory mappings altered to the envelope and synchronization variables. The close range examination power represents a generic main power which the coverage sonification is assisting.

5.4 Transforming Spaces

5.4.1 Motivation

The most recent published work underlying this thesis is Berger (2021a), which like Berger (2020) is tackling the basic problem of self-occluding coverages. But instead of trying to identify a working combination of sensory and spatial mappings for sonification, the idea here is to use interaction to display the data, without requiring a scale change like in Berger and Bill (2019). As described in Section 5.2, the scale change can be an issue even in scanned and simulated environments, once these environments feature many occluding elements or have more complex structures than simple height extrusions on a mostly flat surface

One of the solutions to making coverages visible from an egocentric view that can be found in the literature, are the multi-perspective views from Lorenz et al. (2008) and Veas et al. (2012). In these implementations, the environment around the user is split into an unaltered immediate space, and a far space that is gradually bent upwards to 90 degrees, making data that would usually be out of sight visible on the horizon and in the sky. The rendering of the data for the transformed area is also changed, for example from a 3D city model to a normal top-down 2D cartographic representation.

This of course only works in scenes that are large and flat enough for the result of the bending transformation to remain legible. For more complex environments, the bending needs to be tied to interactions, so that the user can direct and scale the transformations as necessary. How powerful interactions that change the shape and topology of the environment can be can be seen in Bergmann and Lally (2020). Here, a system is shown that allows users to input surface data in the form of (geo-referenced) images, as well as the creation of new geometries. Then, the user can transform the images in various ways: morphing them, cutting them along a line, creating a hole where a polygon was marked, and more. This is done to create new kinds of spatial relations, like straightening a hiking trail around a lake, to get a map representing the viewpoint of someone walking that trail.

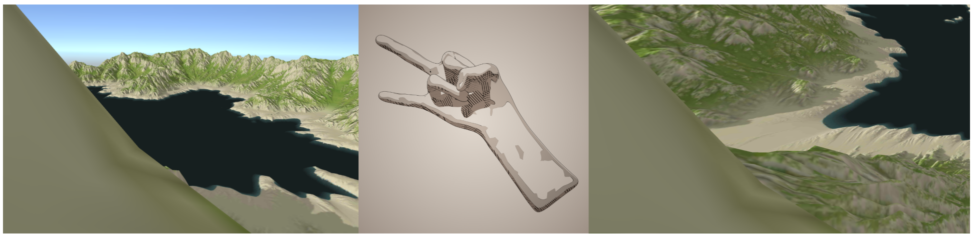
These references were then combined with an idea from Crawford (2019), which contains a collection of body-driven interface prototypes for spatial computing applications. The driving principle behind these interactions was the concept of embodiment and how it could give users entirely new means of utilizing their body in the world. This combination of ideas (Lorenz et al., 2008; Crawford, 2019; Bergmann and Lally, 2020) is ultimately what drove the adoption of the interaction framework from Section 3.3 and the super-power based conception at the end of Chapter 4, which provide the conceptual bridge between many of the concepts that started with the earlier papers.

5.4.2 Technical Summary

The bending operation from Lorenz et al. (2008) can be thought of as a non-affine transformation of the whole virtual environment, i.e. a transformation in which parallel lines do not have to stay parallel (among other factors). Because we want the immediate context of the user to stay unaffected, this transformation needs to happen progressively as distance to the user increases. Both the transformation matrix and the distance factor need to be adjustable depending on the current scene. A user might want to bend towards the sky what is on the other side of a city model, or they might want to straighten out a curved building hallway that is right in front

(A)

(B)



(C)

FIGURE 5.6: Three different emergent hand gestures featured in Berger (2021a): "a) Flattening and stretching. b) A sharp upwards bend. c) A more gradual upwards and sideways bend that brings the obscured part of the coastline into view."

of them. A full description of interaction axes required to control such a non-affine transformation can be found in Berger (2021a), but the final number is 21: 15 matrix values and 6 values for setting a distance cursor.

There is no conventional interface device that is capable of effectively controlling this number of axes. Instead, inspired by Crawford (2019)'s interfaces, we utilize the most capable interaction device a human has and map the 21 values to it: the hand. Specifically, we make use of the concept of microgestures, in which finger joint configurations are used for interaction. Berger (2021a) features an extensive list of all involved joints and their capabilities. The right hand controls the distance cursor as well as teleportation and other interaction state changes. The left hand controls the 15 matrix values. An important aspect of this system is that the mapping of these values to the joints is intentionally kept hidden from the user. The mapping is too complex to usefully explain in a tutorial or legend, so it simply needs to be learned through trial and error. This allows the user to unlock a space of gestural control completely by themselves—the gestures a user becomes familiar with simply emerge during experimentation. Some examples are shown in Figure 5.6.

During implementation, it became clear that it was important to give the scene a certain felt weight. If the transformations were applied too directly, it was almost impossible to process the feedback between the hand state and the visual scene. Once the scene had an appropriate amount of inertia, making it feel like an actual land mass or actual material was shifting and turning, the system became much easier to use. This could be further reinforced by utilizing sound to display to the user how much of an effect the current hand movement has. If the scene has to overcome a lot of inertia and will change completely, a low rumbling sound meant to resemble moving earth is played more and more loudly.

The software itself was created in the Unity engine and makes use of Oculus Quest hand tracking. It makes heavy use of vertex shaders in order to retain a high performance and immediate feedback for the interactions.

5.4.3 Applying the Model

There are some distinct differences this system has to the powers that were already described in this chapter. There is no conventional GIS data being sensified—instead, we are trying to make sense of an environment by manipulating it. Changes to the scene being a direct part of the sensification loop turns this into more of a pragmatic power instead of an epistemic power—we enact control on the environment directly, instead of trying to understand data.

This creates a novel interplay of user interactions "creating" the temporal aspect of the data in the form of an environment that is shifting around us. Figure 5.7 shows how this can be represented in our diagram syntax. Interactions with a scene operand create a feedback in the form of a shapeshifting surface model and moving points that create the aforementioned rumbling sound. The current state of this movement in the scene is directly tied to the current hand configuration of the user, represented by the kinesthetic alignment (H7) variable—even though this is not a technically implemented mapping from a value to sensory variable through some device, the effect exists because of the embodiment itself. As the hand configuration changes, the scene transformation slowly changes with it. In other conceptions of embodiment, this would mean that the hand is embodied by the scene itself.

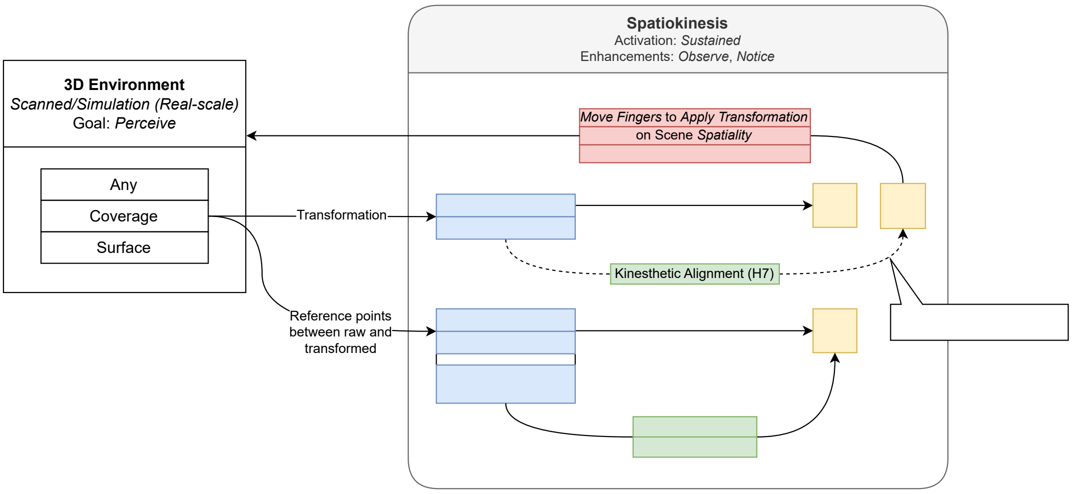


FIGURE 5.7: The model from Chapter 4 applied to the transformation system from Berger (2021a). The kinesthetic alignment is specifically marked, because it is a passive effect of the system and not a true sensory mapping. This example also shows how multiple data sets can be included in one power, that powers can have multiple loops (usually for second order effects that support the primary power), and that there can be interactions that aim at scene operands but are still part of the sensification loop of a larger power.

Such an interface is likely too experimental for deployment in production software, but it shows that being more playful with our interface development is a possibility, especially once embodiment comes into play. Just like we discover objects in the world around us by testing our expectations in how they will react to interactions, an embodied interface can be explored by trying out different *verbs* until the world reacts as we want it to. This discoverability is a common feature of immersive games and could give important hints on how to design analytical systems that are engaging to use.

This interface is also a conclusion of the question of coverage sensification. In the end there are multiple very different approaches to solving this problem, either through strong sensory and spatial mappings, through embodied interactions, or through a combination of the two. Most attemps at designing embodied geosensifications will encounter similar problems—problems that can be solved in many ways. What the model can do is to allow us to think through these different solutions and to integrate different aspects of these systems into one whole. Instead of creating a visualization first and then simply retrofitting a controller-based control scheme, because that is what is expected, thinking in empowerments allows us to integrate senses, data and interactions in a goal-driven fashion.

Chapter 6

Conclusion and Outlook

At the beginning of this thesis, we introduced the concept of embodied geosensifications: Systems which facilitate exploration of and analytical reasoning about geospatial features within a geographic scene, by enabling us to explore and manipulate this data with movements of our body and by displaying aspects of the data to multiple of our sensory modalities.

Four research questions were posed in relation to such systems, the first three of them aimed at individual system aspects: Senses (the different sensory modalities and means of displaying to them), data (the nature of spatiotemporal data in immersive environments) and interactions (the way we utilize our body within these applications). The last question was concerned with how to model embodied geosensifications in a practical way.

After a broad exploration of topics related to cartography, spatial computing and immersive geovisualization, we identified that the current difficulty in creating embodied geosensifications lies not only in hardware development or technical implementation, but in even having complete conceptions of what the design decisions and possible choices within such systems are.

Based on this fact, the answers to the research questions would need to be precisely such conceptions—taxonomies of the individual aspects that drive embodied geosensifications and ways to combine them. Thus, the following research contributions were made throughout this thesis:

For the *sense*-aspect, a full space of multisensory display variables for the geospatial use case was established. These variables were rated for their capabilities in displaying different classes of values, their perceptual qualities, and their technical feasibility. In creating this variable space, a thorough discussion of the aspects of each of the human primary senses in the context of geospatial data sensification was performed.

For the *data*-aspect, the differences between earlier theories of space and the current state of the art in GIS data representations were discussed. Based on the disconnects identified between current GIS data formats and commonly available, highly dynamic real-time data representations, existing theories of spatiotemporal data were adapted into a taxonomy of data types that would be appropriate for use in immersive scenes. The main contribution in this taxonomy is the formal definition of the continuant, a hypothetical data object that is fully dynamic on every one of its aspects and represents an intersection of GIS and real-time data representations. Based on this taxonomy of data types, a secondary taxonomy of their output types was established. This taxonomy describes the basic geometric, temporal and value structures that are available for sensory display once the data arrives in a sensification.

For the *interaction*-aspect, an existing taxonomy of cartographic interaction was extended and adapted into a taxonomy of embodied interactions in geosensifications. The main contribution is the concept of the embodied operator, a description of an act of interaction that includes both the operation that is performed on a scene or data set, and a clear description of the body movement that is necessary to initiate it.

Finally, all three taxonomies were combined into a complete *model* of embodied geosensification. The primary contribution in this model is a diagram language that is constructed from the three taxonomies and several other aspects of the state of the art that are relevant at the intersection points of these taxonomies (like the concept of geographic scenes). Integrated into this diagram language is the other main contribution: the conception of embodied geosensifications as (super-)powers, which can enable clearer design thinking by providing useful metaphors and clear structural units for creating such systems.

The individual steps to constructing this diagram language were first demonstrated for the example of a situated weather prediction system. After the model was complete, it was also applied to three existing research use cases: the sonification of urban traffic noise for visceral data display, the situated sonification of raster data sets, and the interactive embodied exploration of spatially complex 3D environments with VR hand tracking. These three applications, together with the weather predictions example, demonstrate that the model can aid system designers in thinking about possibilities, decisions, issues and constraints in designing these systems, either based on specific issues encountered during implementation or before ever starting implementation.

At the end of this thesis stands a complete diagram language with multiple taxonomies underlying its elements. Using it can promote more integrated thinking, as it lets designers avoid the pitfalls of focusing on only one aspect of such a system (sensification, spatiotemporal data, embodied interactivity) at a time, thus allowing the creation of techniques that from the beginning feature interesting interplay between these aspects.

Future research into embodied geosensifications should use the diagram developed here to find solutions to analytical questions in a variety of practical domains that have intersections with GIS technologies. There also needs to be continued empirical testing of the variables of the multisensory variable space in the context of real applications, as well as research into the advantages of embodied interaction technologies in immersive geographic scenes. Finally, data representations that are working approximations of the continuant data type need to be developed, such that the disconnect between GIS and real-time graphics can be bridged more directly.

Ultimately the hope for the future of such systems is that they become so embodied, so multisensory and seamless, that the technologies enabling them can almost vanish into the background, leaving only the enhancements to human spatial reasoning that they enable. Ideally, instead of just being conceived as metaphorical superpowers, embodied geosensifications will one day turn into true superpowers of spatial reasoning.

Appendix

Spatiotemporal Algebra

The following are the full definitions for the spatiotemporal algebra from Section 3.2, based on a combination of concepts from Goodchild et al. (2007) and Ferreira et al. (2014)'s original algebra. First, the basic elements, which apart from the *Values*-type are considered atomic:

$$
\begin{aligned}
&\textit{type}\ \textbf{Time}\\
&\textit{type}\ \textbf{Instant}\ \text{inherits}\ \textbf{Time}\\
&\textit{type}\ \textbf{Period}\ \text{inherits}\ \textbf{Time}\\
&\\
&\textit{type}\ \textbf{Geo-Object}\\
&\textit{type}\ \textbf{Geo-Field}\\
&\\
&\textit{type}\ \textbf{Numerical}\\
&\textit{type}\ \textbf{Nominal}\\
&\textit{type}\ \textbf{Ordinal}\\
&\\
&\textit{type}\ \textbf{Values}\\
&\textit{operations}:\\
&\quad \text{new}: \{v_1, v_2, ..., v_n\} \rightarrow \textit{Values} \mid v: \textit{Numerical}\ \text{or}\ \textit{Nominal}\ \text{or}\ \textit{Ordinal}
\end{aligned}
\tag{1}
$$

Based on these basic types, we define the full spatiotemporal types:

$$
\begin{aligned}
&\textit{type}\ \textbf{Observations}\ [F: \textit{Type}, P: \textit{Type}, S: \textit{Type}]\\
&\textit{operations}:\\
&\quad \text{new}: \{(F,P,S)_1, (F,P,S)_2, ..., (F,P,S)_n\} \rightarrow \textit{Observations} \mid n > 0\\
&\quad \text{reference}: \textit{Observations} \rightarrow F\\
&\quad \text{positions}: \textit{Observations} \rightarrow \{P_1, ..., P_n\}\\
&\quad \text{sample}: \textit{Observations} \times P \rightarrow S
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
&\textit{type } \textbf{SpatioTemporal}\\
&\textit{operations}:\\
&\quad \text{observations} : SpatioTemporal \rightarrow Observations\\
&\quad \text{begins, ends} : SpatioTemporal \rightarrow Instant\\
&\quad \text{boundary} : SpatioTemporal \rightarrow Geo\text{-}Object\\
&\quad \text{after, before, during} : SpatioTemporal \times Time \rightarrow SpatioTemporal\\
&\quad \text{intersection, difference} : SpatioTemporal \times Geo\text{-}Object \rightarrow \{st_1, ..., st_n\} \mid\\
&\qquad st : SpatioTemporal\\
&\textit{axioms}:\\
&\quad st_1, st_2 : SpatioTemporal;\ t : Time;\ g : Geo\text{-}Object;\\
&\quad \text{before}(st_1, \text{begins}(st_1)) = Null\\
&\quad \text{after}(st_1, \text{ends}(st_1)) = Null\\
&\quad \text{during}(\text{before}(st_1, t), t) = Null\\
&\quad \text{during}(\text{after}(st_1, t), t) = Null\\
&\quad \text{after}(\text{before}(st_1, t), t) = Null\\
&\quad \text{before}(\text{after}(st_1, t), t) = Null\\
&\quad \text{difference}(st_1, \text{boundary}(st_1)) = \varnothing\\
&\quad \text{intersection}(st_1, \text{boundary}(st_1)) = \{st_1\}\\
&\quad \text{within}(\text{boundary}(st_1), g) = TRUE \Rightarrow \text{intersection}(st_1, g) = \{st_1\}\\
&\quad \text{disjoint}(\text{boundary}(st_1), g) = TRUE \Rightarrow \text{intersection}(st_1, g) = \varnothing\\
&\quad st_2 \in \text{intersection}(st_1, g) \Rightarrow \text{difference}(st_2, g) = \varnothing\\
&\quad st_2 \in \text{intersection}(st_1, g) \Rightarrow \text{boundary}(st_2) = g
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
&\textit{type}\ \textbf{TimeSeries}\ [G : \textit{Geo-Object}, T : \textit{Time}, V : \textit{Values}] \\
&\quad \text{inherits}\ \textbf{SpatioTemporal} \\
&\textit{operations} : \\
&\quad \text{new} : \textit{Period} \times \textit{Observations}[G, T, V] \rightarrow \textit{TimeSeries} \\
&\quad \text{values} : \textit{TimeSeries} \times T \rightarrow V \\
&\quad \text{min, max} : \textit{TimeSeries} \rightarrow V \\
&\quad \text{less, greater, equals} : \textit{TimeSeries} \times V \rightarrow \{ts_1, ..., ts_n\} | ts : \textit{TimeSeries} \\
&\textit{axioms} : \\
&\quad ts_1, ts_2 : \textit{TimeSeries};\ t_1, t_n : \textit{Time};\ v : \textit{Values}; \\
&\quad p : \textit{Period};\ obs : \textit{Observations}; \\
&\quad ts_1 = \text{new}(p, obs) \Rightarrow \text{begins}(ts_1) = \text{begin}(p) \\
&\quad ts_1 = \text{new}(p, obs) \Rightarrow \text{ends}(ts_1) = \text{end}(p) \\
&\quad \text{values}(ts_1, t_1) = \text{sample}(\text{observations}(ts_1), t_1) \\
&\quad \text{after}(t_1, \text{ends}(ts_1)) \vee \text{before}(t_1, \text{begins}(ts_1)) \Rightarrow \textit{values}(ts_1, t_1) = \textit{Null} \\
&\quad \text{values}(\text{after}(ts_1, t_1), t_1) = \textit{Null} \\
&\quad \text{values}(\text{before}(ts_1, t_1), t_1) = \textit{Null} \\
&\quad \text{less}(ts_1, \textit{min}(ts_1)) = \varnothing \\
&\quad \text{greater}(ts_1, \textit{max}(ts_1)) = \varnothing \\
&\quad ts_2 \in \text{equals}(ts_1, v) \Rightarrow \text{min}(ts_2) = \text{max}(ts_2) = v \\
&\quad ts_2 \in \text{less}(ts_1, v) \Rightarrow \text{max}(ts_2) < v \\
&\quad ts_2 \in \text{greater}(ts_1, v) \Rightarrow \text{min}(ts_2) > v \\
&\quad \text{boundary}(ts_1) = \text{reference}(\text{observations}(ts_1)) \\
&\quad \text{positions}(\text{observations}(ts_1)) = \{t_1, ..., t_n\} \Rightarrow \textit{begins}(ts_1) \leq t_1 \\
&\quad \text{positions}(\text{observations}(ts_1)) = \{t_1, ..., t_n\} \Rightarrow \textit{ends}(ts_1) \geq t_n
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
&\textit{type } \textbf{Trajectory } [V : \textit{Values}, T : \textit{Time}, G : \textit{Geo-Object}]\\
&\quad \text{inherits } \textbf{SpatioTemporal}\\
&\textit{operations} :\\
&\quad \text{new} : \textit{Period} \times \textit{Observations}[V, T, G] \rightarrow \textit{Trajectory}\\
&\quad \text{geometry} : \textit{Trajectory} \times T \rightarrow G\\
&\textit{axioms} :\\
&\quad tj : \textit{Trajectory};\ t_1, t_n : \textit{Time};\ g : \textit{Geo-Object};\\
&\quad p : \textit{Period};\ obs : \textit{Observations};\\
&\quad tj = \text{new}(p, obs) \Rightarrow \text{begins}(tj) = \text{begin}(p)\\
&\quad tj = \text{new}(p, obs) \Rightarrow \text{ends}(tj) = \text{end}(p)\\
&\quad \text{geometry}(tj, t_1) = \text{sample}(\text{observations}(tj), t_1)\\
&\quad \text{after}(t_1, \text{ends}(tj)) \vee \text{before}(t_1, \text{begins}(tj)) \Rightarrow \textit{geometry}(tj, t_1) = \textit{Null}\\
&\quad \text{geometry}(\text{after}(tj, t_1), t_1) = \textit{Null}\\
&\quad \text{geometry}(\text{before}(tj, t_1), t_1) = \textit{Null}\\
&\quad \text{positions}(\text{observations}(tj)) = \{t_1, ..., t_n\} \Rightarrow \textit{begins}(tj) \leq t_1\\
&\quad \text{positions}(\text{observations}(tj)) = \{t_1, ..., t_n\} \Rightarrow \textit{ends}(tj) \geq t_n\\
&\quad \text{sample}(\text{observations}(tj), t_n) = g \Rightarrow \text{within}(g, \text{boundary}(tj)) = \textit{TRUE}
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
&\textit{type } \textbf{Coverage } [T : Time, G : Geo\text{-}Object, V : Values] \\
&\quad \text{inherits } \textbf{SpatioTemporal} \\
&\textit{operations}: \\
&\quad \text{new} : Geo\text{-}Field \times Observations[T, G, V] \rightarrow Coverage \\
&\quad \text{values} : Coverage \times Geo - Object \rightarrow V \\
&\quad \text{min, max} : Coverage \rightarrow V \\
&\quad \text{less, greater, equals} : Coverage \times V \rightarrow Coverage \\
&\textit{axioms}: \\
&\quad cv_1, cv_2 : Coverage;\ g : Geo\text{-}Object;\ gf : Geo\text{-}Field;\ v : values; \\
&\quad obs : Observations;\ t : Time \\
&\quad cv_1 = \text{new}(gf, obs) \Rightarrow \text{boundary}(cv_1) = g \\
&\quad \text{begins}(cv_1) = \text{begin}(\text{reference}(\text{observations}(cv_1))) \\
&\quad \text{ends}(cv_1) = \text{end}(\text{reference}(\text{observations}(cv_1))) \\
&\quad \text{values}(cv_1, g) = \text{sample}(\text{observations}(cv_1), g) \\
&\quad \text{disjoint}(g, \text{boundary}(cv_1)) = TRUE \Rightarrow values(cv_1, g) = Null \\
&\quad \text{less}(cv_1, \text{min}(cv_1)) = Null \\
&\quad \text{greater}(cv_1, \text{max}(cv_1)) = Null \\
&\quad \text{equals}(cv_1, v) = cv_2 \Rightarrow \text{min}(cv_2) = \text{max}(cv_2) = v \\
&\quad \text{less}(cv_1, v) = cv_2 \Rightarrow \text{max}(cv_2) < v \\
&\quad \text{greater}(cv_1, v) = cv_2 \Rightarrow \text{min}(cv_2) > v \\
&\quad \text{less}(\text{equals}(cv_1, v), v) = Null \\
&\quad \text{greater}(\text{equals}(cv_1, v), v) = Null \\
&\quad cv_2 \in \text{intersection}(cv_1, g) \Rightarrow \text{boundary}(cv_2) = g \\
&\quad cv_2 \in \text{difference}(cv_1, g) \Rightarrow \text{boundary}(cv_2) = \text{boundary}(cv_1)
\end{aligned}
\tag{6}
$$

type **CoverageSeries** $[G : \textit{Geo-Object}, T : \textit{Time}, CV : \textit{Coverage}]$
inherits **SpatioTemporal**
operations :
new : $\textit{Period} \times \textit{Observations}[G, T, CV] \rightarrow \textit{CoverageSeries}$
snapshot : $\textit{CoverageSeries} \times T \rightarrow CV$
timeseries : $\textit{CoverageSeries} \times G \rightarrow \textit{TimeSeries}$
axioms :
$cs : \textit{CoverageSeries};\ c : \textit{Coverage};\ t_1, t_n : \textit{Time};$
$g : \textit{Geo-Object};\ obs : \textit{Observations};\ p : \textit{Period}$
$cs = \text{new}(p, obs) \Rightarrow \text{begins}(cs) = \text{begin}(p)$
$cs = \text{new}(p, obs) \Rightarrow \text{ends}(cs) = \text{end}(p)$
$\text{snapshot}(cs, t_1) = \text{sample}(\textit{observations}(cs), t_1)$ (7)
$\text{snapshot}(\text{after}(cs, t_1), t_1) = \textit{Null}$
$\text{snapshot}(\text{before}(cs, t_1), t_1) = \textit{Null}$
$\text{after}(t_1, \text{ends}(cs)) \vee \text{before}(t_1, \text{begins}(cs)) \Rightarrow \text{snapshot}(cs, t_1) = \textit{Null}$
$\text{begins}(\text{timeseries}(cs, g)) = \text{begins}(cs)$
$\text{ends}(\text{timeseries}(cs, g)) = \text{ends}(cs)$
$\text{boundary}(cs) = \text{reference}(\text{observations}(cs))$
$\text{sample}(\text{observations}(cs), t_1) = c \Rightarrow \text{boundary}(cs) = \text{boundary}(c)$
$\text{sample}(\text{observations}(cs), t_1) = c \Rightarrow \text{begins}(c) = \text{begin}(t_1)$
$\text{sample}(\text{observations}(cs), t_1) = c \Rightarrow \text{ends}(c) = \text{end}(t_1)$
$\text{positions}(\text{observations}(cs)) = \{t_1, ..., t_n\} \Rightarrow \text{begins}(cs) \leq t_1$
$\text{positions}(\text{observations}(cs)) = \{t_1, ..., t_n\} \Rightarrow \text{ends}(cs) \geq t_n$

$$
\begin{aligned}
&\textit{type } \textbf{Continuant } [ID : Values, CS : CoverageSeries, TJ : Trajectory] \\
&\textit{operations} : \\
&\quad \text{new} : ID \times TS \times TJ \rightarrow Continuant \\
&\quad \text{id} : Continuant \rightarrow ID \\
&\quad \text{coverageseries} : Continuant \rightarrow CS \\
&\quad \text{trajectory} : Continuant \rightarrow TJ \\
&\quad \text{state} : Continuant \times Time \rightarrow (Geo\text{-}Object, Coverage) \\
&\textit{axioms} : \\
&\quad c : Continuant;\ t : Time;\ v : Values;\ g : Geo\text{-}Object; \\
&\quad id(c) = \text{reference}(\text{Observations}(\text{trajectory}(c))) \\
&\quad \text{within}(\text{boundary}(\text{trajectory}(c)), \text{boundary}(\text{coverageseries}(c))) = TRUE \\
&\quad \text{begins}(\text{trajectory}(c)) = \text{begins}(\text{coverageseries}(c)) \\
&\quad \text{ends}(\text{trajectory}(c)) = \text{ends}(\text{coverageseries}(C)) \\
&\quad \text{state}(c, t) = (\text{geometry}(\text{trajectory}(c), t), \text{snapshot}(\text{coverageseries}(c), t)) \\
&\quad \text{disjoint}(\text{geometry}(\text{trajectory}(c), t), \\
&\qquad \text{intersection}(g, \text{boundary}(\text{coverageseries}(c)))) \\
&\qquad \rightarrow \text{values}(\text{snapshot}(\text{coverageseries}(c), t), g) = Null
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
&\textit{type } \textbf{Occurrent } [ID : Values, G : Geo\text{-}Object] \\
&\textit{operations} : \\
&\quad \text{new} : ID \times G \times \{con_1, con_2, ..., con_n\} \rightarrow Occurrent \mid \\
&\qquad con : Continuant \text{ and } n \geq 0 \\
&\quad \text{id} : Occurrent \rightarrow ID \\
&\quad \text{time} : G \rightarrow Period \\
&\quad \text{location} : Occurrent \rightarrow G \\
&\quad \text{continuants} : Occurrent \rightarrow \{con_1, con_2, ..., con_n\} \\
&\textit{axioms} : \\
&\quad o : Occurrent;\ c : Continuant;\ t : Time;\ v : Values;\ g : Geo\text{-}Object; \\
&\quad c \in \text{continuants}(o) \wedge \text{time}(\text{location}(o)) = t \Rightarrow \text{sample}(c, t) \neq Null \\
&\quad c \in \text{continuants}(o) \wedge \text{location}(o) = g \\
&\qquad \Rightarrow \text{intersects}(\text{boundary}(\text{trajectory}(c)), g) = TRUE
\end{aligned}
\tag{9}
$$

Bibliography

Ahn, S. J., Bostick, J., Ogle, E., Nowak, K. L., McGillicuddy, K. T., and Bailenson, J. N. (2016). Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *Journal of Computer-Mediated Communication*, 21(6):399–419.

Andrienko, G., Andrienko, N., Dykes, J., Kraak, M.-J., and Schumann, H. (2010). Geova(t)—geospatial visual analytics: Focus on time: Special issue of the international cartographic association commission on geovisualization.

Andrienko, G., Andrienko, N., Schumann, H., and Tominski, C. (2014). Visualization of trajectory attributes in space–time cube and trajectory wall. In *Cartography from Pole to Pole*, pages 157–163. Springer.

Ayoub, A. and Pulijala, Y. (2019). The application of virtual reality and augmented reality in oral & maxillofacial surgery. *BMC Oral Health*, 19(1):1–8.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385.

Beck, J., Rainoldi, M., and Egger, R. (2019). Virtual reality in tourism: A state-of-the-art review. *Tourism Review*, 74(3):586–612.

Berger, M. (2020). Egocentric sonification of continuous spatial data in situated analytics. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 658–659. IEEE.

Berger, M. (2021a). Exploring and transforming spaces through high-dimensional gestural interactions. *Advances in Cartography and GIScience of the ICA*, 3.

Berger, M. (2021b). Modeling multimodal interactions and feedback for embodied geovisualization. In *International Conference on Theory and Application of Diagrams*, pages 488–491. Springer.

Berger, M. and Bill, R. (2019). Combining VR visualization and sonification for immersive exploration of urban noise standards. *Multimodal Technologies and Interaction*, 3(2):34.

Bergmann, L. and Lally, N. (2020). For geographical imagination systems. *Annals of the American Association of Geographers*, pages 1–10.

Bernardes, S., Madden, M., Knight, A., Neel, N., Morgan, N., Cameron, K., and Knox, J. (2018). A multi-component system for data acquisition and visualization in the geosciences based on uavs, augmented and virtuak reality. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Bertin, J. (1967). *Semiology of graphics: Diagrams, networks, maps (translated by William J. Berg)*. University of Wisconsin Press.

Bertin, J. (2000). Matrix theory of graphics. *Information Design Journal*, 10(1):5–19.

Bhatnagar, D. K. (1993). Position trackers for head mounted display systems: A survey. *University of North Carolina, Chapel Hill TR93*, 10.

Bill, R. (2023). *Grundlagen der Geo-Informationssysteme. 7., vollständig neu bearbeitete und erweiterte Auflage.*, volume 7. Wichmann.

Bjork, S. and Redstrom, J. (2000). Redefining the focus and context of focus+context visualization. In *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, pages 85–89. IEEE.

Boeldt, D., McMahon, E., McFaul, M., and Greenleaf, W. (2019). Using virtual reality exposure therapy to enhance treatment of anxiety disorders: Identifying areas of clinical adoption and potential obstacles. *Frontiers in Psychiatry*, 10:773.

Boetje, J. and van Ginkel, S. (2021). The added benefit of an extra practice session in virtual reality on the development of presentation skills: A randomized control trial. *Journal of Computer Assisted Learning*, 37(1):253–264.

Boyandin, I., Bertini, E., Bak, P., and Lalanne, D. (2011). Flowstrates: An approach for visual exploration of temporal origin-destination data. In *Computer Graphics Forum*, volume 30, pages 971–980. Wiley Online Library.

Breslin, P. A. and Spector, A. C. (2008). Mammalian taste perception. *Current Biology*, 18(4):R148–R155.

Bye, K. (2017). An elemental theory of presence + future of AI & interactive storytelling. In *Voices of VR, episode 502*, Podcast audio.

Campbell, A. G., Holz, T., Cosgrove, J., Harlick, M., and O'Sullivan, T. (2019). Uses of virtual reality for communication in financial services: A case study on comparing different telepresence interfaces: Virtual reality compared to video conferencing. In *Future of Information and Communication Conference*, pages 463–481. Springer.

Card, S. (2009). Information visualization. In *Human-Computer Interaction*, pages 199–234. CRC press.

Carpendale, M. S. T. (2003). Considering visual variables as a basis for information visualisation.

Cartwright, S., Samoil, S., Lawton, B., Hu, D., Xie, S., Wang, E., Aminbeidokhti, A., Dawar, S., Dalton, R., Daeijavad, P., et al. (2022). Enhancing reservoir engineering workflows with augmented and virtual reality. In *SPE Canadian Energy Technology Conference*. OnePetro.

Chassin, T., Ingensand, J., Christophe, S., and Touya, G. (2022). Experiencing virtual geographic environment in urban 3d participatory e-planning: A user perspective. *Landscape and Urban Planning*, 224:104432.

Cheiran, J. F., Torres, L. A., da Silva, A. A., Souza, G. A. d., Nedel, L. P., Maciel, A., and Barone, D. A. (2020). Comparing physical and immersive VR prototypes for evaluation of an industrial system user interface. In *Computer Graphics International Conference*, pages 3–15. Springer.

Chen, L., Liang, H.-N., Lu, F., Wang, J., Chen, W., and Yue, Y. (2021). Effect of collaboration mode and position arrangement on immersive analytics tasks in virtual reality: A pilot study. *Applied Sciences*, 11(21):10473.

Chen, M. and Lin, H. (2018). Virtual geographic environments (VGEs): Originating from or beyond virtual reality (VR)? *International Journal of Digital Earth*, 11(4):329–333.

Chen, X., Self, J. Z., House, L., Wenskovitch, J., Sun, M., Wycoff, N., Evia, J. R., and North, C. (2017). Be the data: Embodied visual analytics. *IEEE Transactions on Learning Technologies*, 11(1):81–95.

Cheok, A. D. and Karunanayaka, K. (2018). *Virtual taste and smell technologies for multisensory internet and virtual reality*. Springer.

Chessa, M., Maiello, G., Klein, L. K., Paulun, V. C., and Solari, F. (2019). Grasping objects in immersive virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1749–1754. IEEE.

Chouvardas, V. G., Miliou, A. N., and Hatalis, M. K. (2008). Tactile displays: Overview and recent advances. *Displays*, 29(3):185–194.

Chowdhury, S. and Schnabel, M. A. (2019). Laypeople's collaborative immersive virtual reality design discourse in neighborhood design. *Frontiers in Robotics and AI*, 6:97.

Christou, C., Tzanavari, A., Herakleous, K., and Poullis, C. (2016). Navigation in virtual reality: Comparison of gaze-directed and pointing motion control. In *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, pages 1–6. IEEE.

Çöltekin, A., Bleisch, S., Andrienko, G., and Dykes, J. (2017). Persistent challenges in geovisualization—A community perspective. *International Journal of Cartography*, 3(sup1):115–139.

Çöltekin, A., Lochhead, I., Madden, M., Christophe, S., Devaux, A., Pettit, C., Lock, O., Shukla, S., Herman, L., Stachoň, Z., et al. (2020). Extended reality in spatial sciences: A review of research challenges and future directions. *ISPRS International Journal of Geo-Information*, 9(7):439.

Çöltekin, A., Lokka, I., Zahner, M., and Halounova, L. (2016). On the usability and usefulness of 3d (geo) visualizations—A focus on virtual reality environments. In *XXIII ISPRS Congress, Commission II*. ISPRS.

Cordeil, M., Cunningham, A., Dwyer, T., Thomas, B. H., and Marriott, K. (2017). Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pages 71–83.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In *Theories and methods of spatio-temporal reasoning in geographic space*, pages 65–77. Springer.

Couclelis, H. (1996). Towards an operational typology of geographic entities with ill-defined boundaries. In *Geographic Objects with Indeterminate Boundaries*, pages 45–55. Taylor & Francis.

Crawford, G. (2019). Developing embodied familiarity with hyperphysical phenomena. Master's thesis, Carnegie Mellon University.

Cui, W. (2019). Visual analytics: A comprehensive overview. *IEEE Access*, 7:81555–81573.

Cunningham, A., Hart, J. D., Engelke, U., Adcock, M., and Thomas, B. H. (2021). Towards embodied interaction for geospatial energy sector analytics in immersive environments. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, pages 396–400.

Cutolo, F., Fida, B., Cattari, N., and Ferrari, V. (2019). Software framework for customized augmented reality headsets in medicine. *IEEE Access*, 8:706–720.

Dementyev, A. and Paradiso, J. A. (2014). Wristflex: Low-power gesture input with wrist-worn pressure sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, pages 161–166.

Deng, H., Zhang, L., Mao, X., and Qu, H. (2015). Interactive urban context-aware visualization via multiple disocclusion operators. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1862–1874.

DiBiase, D., MacEachren, A. M., Krygier, J. B., and Reeves, C. (1992). Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems*, 19(4):201–214.

Dinis, F. M., Sanhudo, L., Martins, J. P., and Ramos, N. M. (2020). Improving project communication in the architecture, engineering and construction industry: Coupling virtual reality and laser scanning. *Journal of Building Engineering*, 30:101287.

Dolezal, M., Chmelik, J., and Liarokapis, F. (2020). An immersive virtual environment for collaborative geovisualization. *arXiv preprint arXiv:2010.06279*.

Dos Santos, S. and Brodlie, K. (2004). Gaining understanding of multivariate and multidimensional data through visualization. *Computers & Graphics*, 28(3):311–325.

Dübel, S., Röhlig, M., Tominski, C., and Schumann, H. (2017). Visualizing 3d terrain, geo-spatial data, and uncertainty. In *Informatics*, volume 4, page 6. MDPI.

Dubus, G. and Bresin, R. (2013). A systematic review of mapping strategies for the sonification of physical quantities. *PloS One*, 8(12):e82491.

Dudley, J., Benko, H., Wigdor, D., and Kristensson, P. O. (2019). Performance envelopes of virtual keyboard text input strategies in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 289–300. IEEE.

Dwyer, T., Cordeil, M., Czauderna, T., Haghighi, P. D., Ens, B., Goodwin, S., Jenny, B., Marriott, K., and Wybrow, M. (2020). The data visualisation and immersive analytics research lab at monash university. *Visual Informatics*, 4(4):41–49.

Edler, D., Kühne, O., Keil, J., and Dickmann, F. (2019). Audiovisual cartography: Established and new multimedia approaches to represent soundscapes. *KN-Journal of Cartography and Geographic Information*, 69(1):5–17.

Edwards, M. (2019). *Fragrances of the world, 33rd edition*. Fragrances of the World.

Engeln, L., Hube, N., and Groh, R. (2018). Immersive VisualAudioDesign: Spectral editing in VR. In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*, pages 1–4. Association for Computing Machinery.

Ens, B., Cordeil, M., North, C., Dwyer, T., Besançon, L., Prouzeau, A., Liu, J., Cunningham, A., Drogemuller, A., Satriadi, K. A., et al. (2022). Immersive analytics 2.0: Spatial and embodied sensemaking. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–7.

Feng, Y. and Mueller, B. (2019). The state of augmented reality advertising around the globe: A multi-cultural content analysis. *Journal of Promotion Management*, 25(4):453–475.

Ferreira, K. R., Camara, G., and Monteiro, A. M. V. (2014). An algebra for spatiotemporal data: From observations to events. *Transactions in GIS*, 18(2):253–269.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8(5):347–358.

Frisoli, A., Rocchi, F., Marcheschi, S., Dettori, A., Salsedo, F., and Bergamasco, M. (2005). A new force-feedback arm exoskeleton for haptic interaction in virtual environments. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator systems. World Haptics Conference*, pages 195–201. IEEE.

Gaitatzes, A., Christopoulos, D., and Roussou, M. (2001). Reviving the past: Cultural heritage meets virtual reality. In *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage*, pages 103–110.

Galton, A. (2008). Experience and history: Processes and their relation to events. *Journal of Logic and Computation*, 18(3):323–340.

Garcia-Ruiz, M. A., Kapralos, B., and Rebolledo-Mendez, G. (2021). An overview of olfactory displays in education and training. *Multimodal Technologies and Interaction*, 5(10):64.

Garlandini, S. and Fabrikant, S. I. (2009). Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In *International Conference on Spatial Information Theory*, pages 195–211. Springer.

Gaydos, D. A., Petrasova, A., Cobb, R. C., and Meentemeyer, R. K. (2019). Forecasting and control of emerging infectious forest disease through participatory modelling. *Philosophical Transactions of the Royal Society B*, 374(1776):20180283.

Globa, A., Beza, B. B., and Wang, R. (2022). Towards multi-sensory design: Placemaking through immersive environments-evaluation of the approach. *Expert Systems with Applications*, page 117614.

Goodchild, M. F. (1989). Modeling error in objects and fields. *Accuracy of Spatial Databases*, 113.

Goodchild, M. F., Yuan, M., and Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3):239–260.

Gračanin, D. (2018). Immersion versus embodiment: Embodied cognition for immersive analytics in mixed reality environments. In *International Conference on Augmented Cognition*, pages 355–368. Springer.

Greenhalgh, C. and Benford, S. (1995). Massive: A collaborative virtual environment for teleconferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3):239–261.

Greenwold, S. (2003). Spatial computing. *Massachusetts Institute of Technology, Master*.

Griffin, A. L. (2001). Feeling it out: The use of haptic visualization for exploratory geographic analysis. *Cartographic Perspectives*, pages 12–29.

Griffin, A. L., Robinson, A. C., and Roth, R. E. (2017). Envisioning the future of cartographic research.

Grubert, J., Langlotz, T., Zollmann, S., and Regenbrecht, H. (2016). Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1706–1724.

Guarese, R., Becker, J., Fensterseifer, H., Walter, M., Freitas, C., Nedel, L., and Maciel, A. (2020). Augmented situated visualization for spatial and context-aware decision-making. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 1–5.

Haber, R. B. and McNabb, D. A. (1990). Visualization idioms: A conceptual model for scientific visualization systems. *Visualization in Scientific Computing*, 74:93.

Hamburger, K. and Knauff, M. (2019). Odors can serve as landmarks in human wayfinding. *Cognitive Science*, 43(11):e12798.

Harding, C., Kakadiaris, I. A., Casey, J. F., and Loftin, R. B. (2002). A multi-sensory system for the investigation of geoscientific data. *Computers & Graphics*, 26(2):259–269.

Harding, C. and Souleyrette, R. R. (2009). Investigating the use of 3d graphics, haptics (touch), and sound for highway location planning. *Computer-Aided Civil and Infrastructure Engineering*, 1(25):20–38.

Harmon, B. A. (2016). Embodied spatial thinking in tangible computing. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 693–696.

Hartson, R. and Pyla, P. S. (2018). *The UX book: Agile UX design for a quality user experience*. Morgan Kaufmann.

Harvey, F. and Losang, E. (2019). Bertin's matrix concepts reconsidered: transformations of semantics and semiotics to support geovisualization use. *Cartography and Geographic Information Science*, 46(2):152–162.

Hasan, S. M., Lee, K., Moon, D., Kwon, S., Jinwoo, S., and Lee, S. (2022). Augmented reality and digital twin system for interaction with construction machinery. *Journal of Asian Architecture and Building Engineering*, 21(2):564–574.

Hasegawa, K., Qiu, L., and Shinoda, H. (2018). Midair ultrasound fragrance rendering. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1477–1485.

Horbiński, T. and Zagata, K. (2022). Interpretation of map symbols in the context of gamers' age and experience. *ISPRS International Journal of Geo-Information*, 11(2):150.

Hoshi, T., Takahashi, M., Iwamoto, T., and Shinoda, H. (2010). Noncontact tactile display based on radiation pressure of airborne ultrasound. *IEEE Transactions on Haptics*, 3(3):155–165.

Howard, M. C., Gutworth, M. B., and Jacobs, R. R. (2021). A meta-analysis of virtual reality training programs. *Computers in Human Behavior*, 121:106808.

Hruby, F. (2019a). From third-person to first-person cartographies with immersive virtual environments. *Proceedings of the International Cartographic Association*, 2:44.

Hruby, F. (2019b). The sound of being there: Audiovisual cartography with immersive virtual environments. *KN-Journal of Cartography and Geographic Information*, 69(1):19–28.

Hruby, F., Castellanos, I., and Ressl, R. (2021). Cartographic scale in immersive virtual environments. *KN-Journal of Cartography and Geographic Information*, 71(1):45–51.

Hruby, F., Ressl, R., Borbolla del Valle, G. D. l., Rodríguez Paz, O., Aguilar Sierra, V., and Coutiño, J. H. M. (2019a). Real geographies in virtual space: A practical workflow for geovisualization with immersive vr. In *International Congress of Telematics and Computing*, pages 3–15. Springer.

Hruby, F., Ressl, R., and de la Borbolla Del Valle, G. (2019b). Geovisualization with immersive virtual environments in theory and practice. *International Journal of Digital Earth*, 12(2):123–136.

Hruby, F., Sánchez, L. F. Á., Ressl, R., and Escobar-Briones, E. G. (2020). An empirical study on spatial presence in immersive geo-environments. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(2):155–163.

Hu, C., Li, J., Xiao, C., Wang, K., and Chen, N. (2020). Soco-field: Observation capability representation for geotask-oriented multi-sensor planning cognition. *International Journal of Geographical Information Science*, 34(2):205–228.

Huang, E. Y., Knight, S., Guetter, C. R., Davis, C. H., Moller, M., Slama, E., and Crandall, M. (2019). Telemedicine and telementoring in the surgical specialties: A narrative review. *The American Journal of Surgery*, 218(4):760–766.

Hugues, O., Cieutat, J.-M., and Guitton, P. (2011). GIS and augmented reality: State of the art and issues. *Handbook of Augmented Reality*, pages 721–740.

Imran, E., Adanir, N., and Khurshid, Z. (2021). Significance of haptic and virtual reality simulation (VRS) in the dental education: A review of literature. *Applied Sciences*, 11(21):10196.

Ishizuka, H. and Miki, N. (2015). MEMS-based tactile displays. *Displays*, 37:25–32.

J. W. Thrower, N. (2007). *Maps and Civilization. Cartography in Culture and Society, Third Edition.* University of Chicago Press.

Jackson, R. L. and Fagan, E. (2000). Collaboration and learning within immersive virtual reality. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*, pages 83–92.

Jensen, C. G. (2017). Collaboration and dialogue in virtual reality. *Journal of Problem Based Learning in Higher Education*, 5(1).

Jochecová, K., Černỳ, M., Stachoň, Z., Švedová, H., Káčová, N., Chmelík, J., Brůža, V., Kvarda, O., Ugwitz, P., Šašinková, A., et al. (2022). Geography education in a collaborative virtual environment: A qualitative study on geography teachers. *ISPRS International Journal of Geo-Information*, 11(3):180.

Kaeppler, K. and Mueller, F. (2013). Odor classification: A review of factors influencing perception-based odor arrangements. *Chemical senses*, 38(3):189–209.

Kamilaris, A. and Ostermann, F. O. (2018). Geospatial analysis and the internet of things. *ISPRS International Journal of Geo-information*, 7(7):269.

Kapralos, B., Collins, K., and Uribe-Quevedo, A. (2017). The senses and virtual environments. *The Senses and Society*, 12(1):69–75.

Kapur, A., Kapur, S., and Maes, P. (2018). Alterego: A personalized wearable silent speech interface. In *23rd International Conference on Intelligent User Interfaces*, pages 43–53.

Käser, D. P., Parker, E., Glazier, A., Podwal, M., Seegmiller, M., Wang, C.-P., Karlsson, P., Ashkenazi, N., Kim, J., Le, A., et al. (2017). The making of Google Earth VR. In *ACM SIGGRAPH 2017 Talks*, page 63.

Keil, J., Edler, D., Schmitt, T., and Dickmann, F. (2021). Creating immersive virtual environments based on open geospatial data and game engines. *KN-Journal of Cartography and Geographic Information*, 71(1):53–65.

Keim, D. A., Mansmann, F., Schneidewind, J., and Ziegler, H. (2006). Challenges in visual data analysis. In *Tenth International Conference on Information Visualisation (IV'06)*, pages 9–16. IEEE.

Kelling, S. T. and Halpern, B. P. (1988). Taste judgements and gustatory stimulus duration: Taste quality, taste intensity, and reaction time. *Chemical Senses*, 13(4):559–586.

Kephalopoulos, S., Paviotti, M., and Ledee, F. A. (2012). Common noise assessment methods in europe (CNOSSOS-EU).

Khanal, P., Vankipuram, A., Ashby, A., Vankipuram, M., Gupta, A., Drumm-Gurnee, D., Josey, K., Tinker, L., and Smith, M. (2014). Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles. *Journal of Biomedical Informatics*, 51:49–59.

Kim, J. (2021). Advertising in the metaverse: Research agenda. *Journal of Interactive Advertising*, 21(3):141–144.

Kjenstad, K. (2006). On the integration of object-based models and field-based models in GIS. *International Journal of Geographical Information Science*, 20(05):491–509.

Klaas, E. and Roopaei, M. (2021). Immersive analytics application in smart agriculture and animal behavior. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0290–0296. IEEE.

Kluge, S., Gladisch, S., von Lukas, U. F., Staadt, O., and Tominski, C. (2019). Virtual lenses as embodied tools for immersive analytics. *arXiv preprint arXiv:1911.10044*.

Köbben, B. and Yaman, M. (1995). Evaluating dynamic visual variables. In *Proceedings of the Seminar on Teaching Animated Cartography, Madrid, Spain*, pages 45–51.

Kockro, R. A., Stadie, A., Schwandt, E., Reisch, R., Charalampaki, C., Ng, I., Yeo, T. T., Hwang, P., Serra, L., and Perneczky, A. (2007). A collaborative virtual reality environment for neurosurgical planning and training. *Operative Neurosurgery*, 61(suppl_5):ONSE379–ONSE391.

Kreuziger, U., Hehl, K., and Möller, M. (2015). Optimierungsansätze zur Positionierung und Registrierung geographischer Augmented-Reality-Systeme. *AGIT: Journal für Angewandte Geoinformatik*, 1:486–495.

Krüger, R., Thom, D., Wörner, M., Bosch, H., and Ertl, T. (2013). Trajectorylenses—A set-based filtering and exploration technique for long-term trajectory data. In *Computer Graphics Forum*, volume 32, pages 451–460. Wiley Online Library.

Krygier, J. B. (1994). Sound and geographic visualization. In *Modern Cartography Series*, volume 2, pages 149–166. Elsevier.

Kühne, O. and Edler, D. (2018). Multisensorische Landschaften—Die Bedeutung des Nicht-Visuellen bei der sozialen und individuellen Konstruktion von Landschaft und Herausforderungen für ihre Erfassung und Wiedergabe. *Berichte. Geographie und Landeskunde*, 92(1):27–45.

Lammert-Siepmann, N., Bestgen, A.-K., Edler, D., Kuchinke, L., and Dickmann, F. (2017). Audiovisual communication of object-names improves the spatial accuracy of recalled object-locations in topographic maps. *PLoS One*, 12(10):e0186065.

Langlotz, T., Nguyen, T., Schmalstieg, D., and Grasset, R. (2014). Next-generation augmented reality browsers: Rich, seamless, and adaptive. *Proceedings of the IEEE*, 102(2):155–169.

Lau, C. W., Qu, Z., Draper, D., Quan, R., Braytee, A., Bluff, A., Zhang, D., Johnston, A., Kennedy, P. J., Simoff, S., et al. (2022). Virtual reality for the observation of oncology models (VROOM): Immersive analytics for oncology patient cohorts. *Scientific Reports*, 12.

Lee, B., Hu, X., Cordeil, M., Prouzeau, A., Jenny, B., and Dwyer, T. (2020). Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1171–1181.

Lee, J., Lee, H., Eizad, A., and Yoon, J. (2022). A pilot study on perception of direction cues delivered using a portable electro-tactile biofeedback device. In *2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 01–05. IEEE.

Lee-Cultura, S. (2019). Tetrotation: Utilising multimodal analytics and gestural interaction to nurture mental rotation skills. In *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 715–718.

Li, W., Zhu, J., Gong, Y., Zhu, Q., Xu, B., and Chen, M. (2022). An optimal selection method for debris flow scene symbols considering public cognition differences. *International Journal of Disaster Risk Reduction*, 68:102698.

Lin, H., Chen, M., Lu, G., Zhu, Q., Gong, J., You, X., Wen, Y., Xu, B., and Hu, M. (2013). Virtual geographic environments (VGEs): A new generation of geographic analysis tool. *Earth-Science Reviews*, 126:74–84.

Lin, H., Xu, B., Chen, Y., Li, W., You, L., and He, J. (2022). VGEs as a new platform for urban modeling and simulation. *Sustainability*, 14(13):7980.

Lindemann, B. (2001). Receptors and transduction in taste. *Nature*, 413(6852):219–225.

Liu, Q., Pan, J., and Li, G. (2020). Research of immersive geospatial data visualization based on embodied cognition. In *2020 International Conference on Innovation Design and Digital Technology (ICIDDT)*, pages 358–362. IEEE.

Lorenz, H., Trapp, M., Döllner, J., and Jobst, M. (2008). Interactive multi-perspective views of virtual 3d landscape and city models. In *The European Information Society*, pages 301–321. Springer.

Lou, B. (2012). *The science of sense*. Department of Mathematics, Tongji University, Shanghai, China.

Lü, G., Chen, M., Yuan, L., Zhou, L., Wen, Y., Wu, M., Hu, B., Yu, Z., Yue, S., and Sheng, Y. (2018). Geographic scenario: A possible foundation for further development of virtual geographic environments. *International Journal of Digital Earth*, 11(4):356–368.

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design*. Guilford Press.

MacEachren, A. M., Edsall, R., Haug, D., Baxter, R., Otto, G., Masters, R., Fuhrmann, S., and Qian, L. (1999). Virtual environments for geographic visualization: Potential and challenges. In *Proceedings of the 1999 Workshop on New Paradigms in Information Visualization and Manipulation in Conjunction with the Eighth ACM Internation Conference on Information and Knowledge Management*, pages 35–40.

Mahmood, T., Butler, E., Davis, N., Huang, J., and Lu, A. (2018). Building multiple coordinated spaces for effective immersive analytics through distributed cognition. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–11. IEEE.

Maiese, M. (2018). Can the mind be embodied, enactive, affective, and extended? *Phenomenology and the Cognitive Sciences*, 17(2):343–361.

Marriott, K., Schreiber, F., Dwyer, T., Klein, K., Riche, N. H., Itoh, T., Stuerzlinger, W., and Thomas, B. H. (2018). *Immersive analytics*, volume 11190. Springer.

Matamala-Gomez, M., Bottiroli, S., Realdon, O., Riva, G., Galvagni, L., Platz, T., Sandrini, G., De Icco, R., and Tassorelli, C. (2021). Telemedicine and virtual reality at time of covid-19 pandemic: An overview for future perspectives in neurorehabilitation. *Frontiers in Neurology*, 12:646902.

McCormack, J., Roberts, J. C., Bach, B., Freitas, C. D. S., Itoh, T., Hurter, C., and Marriott, K. (2018). Multisensory immersive analytics. In *Immersive Analytics*, pages 57–94. Springer.

Medellin-Castillo, H. I., Zaragoza-Siqueiros, J., Govea-Valladares, E. H., de la Garza-Camargo, H., Lim, T., and Ritchie, J. M. (2021). Haptic-enabled virtual training in orthognathic surgery. *Virtual Reality*, 25(1):53–67.

Meier, S., Heidmann, F., and Thom, A. (2014). Heattile, a new method for heatmap implementations for mobile web-based cartographic applications. In *Thematic Cartography for the Society*, pages 33–44. Springer.

Meng, L. (2020). An ieee value loop of human-technology collaboration in geospatial information science. *Geo-spatial Information Science*, 23(1):61–67.

Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Cifuentes, L., Kwok, O.-m., and Davis, T. J. (2013). Exploring 3-d virtual reality technology for spatial ability and chemistry achievement. *Journal of Computer Assisted Learning*, 29(6):579–590.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106.

Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329.

Millar, G. C., Tabrizian, P., Petrasova, A., Petras, V., Harmon, B., Mitasova, H., and Meentemeyer, R. K. (2018). Tangible landscape: A hands-on method for teaching terrain analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Miller, K. E. and Muth, E. R. (2004). Efficacy of acupressure and acustimulation bands for the prevention of motion sickness. *Aviation, Space, and Environmental Medicine*, 75(3):227–234.

Moloney, J., Spehar, B., Globa, A., and Wang, R. (2018). The affordance of virtual reality to enable the sensory representation of multi-dimensional data for immersive analytics: From experience to insight. *Journal of Big Data*, 5(1):1–19.

Mori, S., Ikeda, S., and Saito, H. (2017). A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects. *IPSJ Transactions on Computer Vision and Applications*, 9(1):1–14.

Morrison, J. L. (1974). A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *International Yearbook of Cartography*, 14(1974):115–27.

Mueller, F., Dwyer, T., Goodwin, S., Marriott, K., Deng, J., D. Phan, H., Lin, J., Chen, K.-T., Wang, Y., and Ashok Khot, R. (2021). Data as delight: Eating data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2(1):486–497.

Nakamoto, T., Kinoshita, M., Murakami, K., and Yossiri, A. (2009). Demonstration of improved olfactory display using rapidly-switching solenoid valves. In *2009 IEEE Virtual Reality Conference*, pages 301–302. IEEE.

Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M. (2011). Meta cookie+: An illusion-based gustatory display. In *Virtual and Mixed Reality-New Trends: International Conference, Virtual and Mixed Reality 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part I 4*, pages 260–269. Springer.

Newbury, R., Satriadi, K. A., Bolton, J., Liu, J., Cordeil, M., Prouzeau, A., and Jenny, B. (2021). Embodied gesture interaction for immersive maps. *Cartography and Geographic Information Science*, 48(5):417–431.

Newell, R., Canessa, R., and Sharma, T. (2017). Visualizing our options for coastal places: Exploring realistic immersive geovisualizations as tools for inclusive approaches to coastal planning and management. *Frontiers in Marine Science*, 4:290.

Newen, A., De Bruin, L., and Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. Oxford University Press.

Ogi, T. and Hirose, M. (1997). Usage of multisensory information in scientific data sensualization. *Multimedia Systems*, 5(2):86–92.

O'Banion, M. S., Lewis, N. S., Boyce, M. W., Laughlin, J., and Majkowicz, D. C. (2022). Use of an augmented reality sand table for satellite remote sensing education. *Journal of Geography in Higher Education*, pages 1–12.

Pędzich, P. (2019). A low distortion conformal projection of a tri-axial ellipsoid and its application for mapping of extra-terrestrial objects. *Planetary and Space Science*, 178:104697.

Perkins, C. (2009). Performative and embodied mapping. *International Encyclopedia of Human Geography*, 8:126–32.

Petrasova, A., Harmon, B., Petras, V., Tabrizian, P., and Mitasova, H. (2018). *Tangible modeling with open source GIS*. Springer.

Peuquet, D. J. (1988). Representations of geographic space: Toward a conceptual synthesis. *Annals of the Association of American Geographers*, 78(3):375–394.

Philips, A., Walz, A., Bergner, A., Graeff, T., Heistermann, M., Kienzler, S., Korup, O., Lipp, T., Schwanghart, W., and Zeilinger, G. (2015). Immersive 3d geovisualization in higher education. *Journal of Geography in Higher Education*, 39(3):437–449.

Postert, P., Berger, M., and Bill, R. (2021). Utilizing CityGML for AR-labeling and occlusion in urban spaces. In *Advances and New Trends in Environmental Informatics*, pages 209–223. Springer.

Prouzeau, A., Cordeil, M., Robin, C., Ens, B., Thomas, B. H., and Dwyer, T. (2019). Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3d scatterplots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Rajguru, C., Obrist, M., and Memoli, G. (2020). Spatial soundscapes and virtual worlds: Challenges and opportunities. *Frontiers in Psychology*, 11:569056.

Rao, S. M., Mayer, A. R., and Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, 4(3):317–323.

Rice, M., Jacobson, R. D., Golledge, R. G., and Jones, D. (2005). Design considerations for haptic and auditory map interfaces. *Cartography and Geographic Information Science*, 32(4):381–391.

Rink, K., Chen, C., Bilke, L., Liao, Z., Rinke, K., Frassl, M., Yue, T., and Kolditz, O. (2018). Virtual geographic environments for water pollution control. *International Journal of Digital Earth*, 11(4):397–407.

Roberts, D. and Roberts, N. J. (2014). Maximising sensory learning through immersive education. *Journal of Nursing Education and Practice*, 4(10):74–79.

Roberts, J. C. (2007). State of the art: Coordinated & multiple views in exploratory visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, pages 61–71. IEEE.

Roberts, J. C. and Walker, R. (2010). Using all our senses: The need for a unified theoretical approach to multi-sensory information visualization. In *IEEE VisWeek 2010 Workshop: The Role of Theory in Information Visualization, Salt Lake City, Utah*.

Rogers, M. P., DeSantis, A. J., Janjua, H., Barry, T. M., and Kuo, P. C. (2021). The future surgical training paradigm: Virtual reality and machine learning in surgical education. *Surgery*, 169(5):1250–1252.

Röhlig, M., Luboschik, M., and Schumann, H. (2017). Visibility widgets for unveiling occluded data in 3d terrain visualization. *Journal of Visual Languages & Computing*, 42:86–98.

Roth, R. E. (2013a). Cartographic interaction: What we know and what we need to know. *The Journal of Spatial Information Science*, 6(6):59–115.

Roth, R. E. (2013b). An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2356–2365.

Roth, R. E. (2017). Visual variables. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, pages 1–11.

Roth, R. E., Çöltekin, A., Delazari, L., Filho, H. F., Griffin, A., Hall, A., Korpi, J., Lokka, I., Mendonça, A., Ooms, K., et al. (2017). User studies in cartography: Opportunities for empirical research on interactive maps and visualizations. *International Journal of Cartography*, 3(sup1):61–89.

Rydvanskiy, R. and Hedley, N. (2020). 3d geovisualization interfaces as flood risk management platforms: Capability, potential, and implications for practice. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 55(4):281–290.

Rydvanskiy, R. and Hedley, N. (2021). Mixed reality flood visualizations: Reflections on development and usability of current systems. *ISPRS International Journal of Geo-Information*, 10(2):82.

Rylov, M. A. and Reimer, A. W. (2014). A comprehensive multi-criteria model for high cartographic quality point-feature label placement. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 49(1):52–68.

Rzeszewski, M. and Naji, J. (2022). Literary placemaking and narrative immersion in extended reality virtual geographic environments. *International Journal of Digital Earth*, 15(1):853–867.

Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., and Keim, D. A. (2014). Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604–1613.

Safadel, P. and White, D. (2020). Effectiveness of computer-generated virtual reality (VR) in learning and teaching environments with spatial frameworks. *Applied Sciences*, 10(16):5438.

Safikhani, S., Keller, S., Schweiger, G., and Pirker, J. (2022). Immersive virtual reality for extending the potential of building information modeling in architecture, engineering, and construction sector: Systematic review. *International Journal of Digital Earth*, 15(1):503–526.

Saga, S. (2015). Heathapt thermal radiation-based haptic display. *Haptic Interaction: Perception, Devices and Applications*, pages 105–107.

Sardana, D., Kahu, S. Y., Gračanin, D., and Matković, K. (2021). Multi-modal data exploration in a mixed reality environment using coordinated multiple views. In *International Conference on Human-Computer Interaction*, pages 337–356. Springer.

Schetinger, V., Pérez-Messina, I., Guarese, R., and Filipov, V. (2021). Xenakis: Experimenting with data, cities, and sounds. *arXiv preprint arXiv:2109.14992*.

Schito, J. and Fabrikant, S. I. (2018). Exploring maps by sounds: Using parameter mapping sonification to make digital elevation models audible. *International Journal of Geographical Information Science*, 32(5):874–906.

Schneider, A. and Feussner, H. (2017). Chapter 10—Mechatronic support systems and robots. In Schneider, A. and Feussner, H., editors, *Biomedical Engineering in Gastrointestinal Surgery*, pages 387–441. Academic Press.

Schumann, H. and Müller, W. (2013). *Visualisierung: Grundlagen und allgemeine Methoden*. Springer-Verlag.

Schwarz, M. and Hamburger, K. (2022). Modality switching in landmark-based wayfinding. *Frontiers in Psychology*, 13.

Shams, L. and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11):411–417.

Shapiro, L. (2010). *Embodied cognition*. Routledge.

Sheidin, J., Lanir, J., Bak, P., and Kuflik, T. (2017). Time-ray maps: Visualization of spatial and temporal evolution of news stories. In *EuroVis (Short Papers)*, pages 85–89.

Shen, Y., Ong, S. K., and Nee, A. Y. (2010). Augmented reality for collaborative product design and development. *Design Studies*, 31(2):118–145.

Shneiderman, B. (1994). Dynamic queries for visual information seeking. *IEEE Software*, 11(6):70–77.

Siepmann, N., Edler, D., Keil, J., Kuchinke, L., and Dickmann, F. (2020). The position of sound in audiovisual maps: An experimental study of performance in spatial memory. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 55(2):136–150.

Simpson, D. et al. (1974). The choice of control system for the multimovement prosthesis: Extended physiological proprioception (EPP). *The Control of Upper-Extremity Prostheses and Orthoses*, pages 146–150.

Sinton, D. (1978). The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. *Harvard Papers on Geographic Information Systems*.

Slater, M. (2007). The concept of presence and its measurement. *PEACH Summer School, Santorini, Greece*.

Slater, M. and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616.

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., and Hedley, N. R. (2001). Cognitive and usability issues in geovisualization. *Cartography and Geographic Information Science*, 28(1):61–75.

Song, Y., Gong, J., Li, Y., Cui, T., Fang, L., and Cao, W. (2013). Crowd evacuation simulation for bioterrorism in micro-spatial environments based on virtual geographic environments. *Safety Science*, 53:105–113.

Souza, V., Maciel, A., Nedel, L., Kopper, R., Loges, K., and Schlemmer, E. (2020). The effect of virtual reality on knowledge transfer and retention in collaborative group-based learning for neuroanatomy students. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, pages 92–101. IEEE.

Spence, C. (2019). On the relationship(s) between color and taste/flavor. *Experimental Psychology*.

Spur, M., Tourre, V., David, E., Moreau, G., and Le Callet, P. (2020). Mapstack: Exploring multilayered geospatial data in virtual reality. In *11th International Conference on Information Visualization Theory and Applications*, pages 88–99. SCITEPRESS-Science and Technology Publications.

Spur, M., Tourre, V., Moreau, G., and Le Callet, P. (2022). Virtual data sphere: Inverse stereographic projection for immersive multi-perspective geovisualization. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4:235–242.

Ssin, S. Y., Walsh, J. A., Smith, R. T., Cunningham, A., and Thomas, B. H. (2019). Geogate: Correlating geo-temporal datasets using an augmented reality space-time cube and tangible interactions. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 210–219. IEEE.

Steinicke, F., Bruder, G., Jerald, J., Frenz, H., and Lappe, M. (2008). Analyses of human sensitivity to redirected walking. In *Proceedings of the 2008 ACM symposium on Virtual Reality Software and Technology*, pages 149–156.

Strugnell, C. and Jones, L. (1999). Consumer perceptions and opinions of fragrances in household products. *Nutrition & Food Science*, 99(4).

Sugimoto, S., Noguchi, D., Bannnai, Y., and Okada, K. (2010). Ink jet olfactory display enabling instantaneous switches of scents. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 301–310.

Sutherland, I. E. et al. (1965). The ultimate display. In *Proceedings of the IFIP Congress*, volume 2, pages 506–508. New York.

Tabrizian, P., Petrasova, A., Harmon, B., Petras, V., Mitasova, H., and Meentemeyer, R. (2016). Immersive tangible geospatial modeling. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 1–4.

Tak, S. and Toet, A. (2013). Towards interactive multisensory data representations. In *GRAPP/IVAPP*, pages 558–561.

Taketomi, T., Uchiyama, H., and Ikeda, S. (2017). Visual slam algorithms: A survey from 2010 to 2016. *IPSJ Transactions on Computer Vision and Applications*, 9(1):1–11.

Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K., and Li, J. (2020). Toronto-3D: A large-scale mobile lidar dataset for semantic segmentation of urban roadways. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 202–203.

Tanaka, Y., Yamauchi, H., and Amemiya, K. (2002). Wearable haptic display for immersive virtual environment. In *Proceedings of the JFPS International Symposium on Fluid Power*, volume 5-2, pages 309–314. The Japan Fluid Power System Society.

Tominski, C., Abello, J., Van Ham, F., and Schumann, H. (2006). Fisheye tree views and lenses for graph visualization. In *Tenth International Conference on Information Visualisation (IV'06)*, pages 17–24. IEEE.

Tominski, C., Schumann, H., Andrienko, G., and Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2565–2574.

van Ginkel, S., Gulikers, J., Biemans, H., Noroozi, O., Roozen, M., Bos, T., van Tilborg, R., van Halteren, M., and Mulder, M. (2019). Fostering oral presentation competence through a virtual reality-based task for delivering feedback. *Computers & Education*, 134:78–97.

Vasconcellos, R. (1991). Knowing the amazon through tactual graphics. In *Proceedings 15th International*.

Veas, E., Grasset, R., Kruijff, E., and Schmalstieg, D. (2012). Extended overview techniques for outdoor augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):565–572.

Wagner Filho, J. A., Stuerzlinger, W., and Nedel, L. (2019). Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):514–524.

Wan, J., Zheng, Y., Li, Y., Mei, H., Lin, L., and Kuang, L. (2020). Oil depot safety inspection and emergency training system based on virtual reality technology. In *IOP Conference Series: Materials Science and Engineering*, volume 782. IOP Publishing.

Wang, Z., Li, B., Hedgpeth, T., and Haven, T. (2009). Instant tactile-audio map: Enabling access to digital maps for people with visual impairment. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 43–50.

Wesson, D. W. and Wilson, D. A. (2010). Smelling sounds: Olfactory–auditory sensory convergence in the olfactory tubercle. *Journal of Neuroscience*, 30(8):3013–3021.

Westphal, S. and Berger, M. (2020). Entwicklung einer Virtual Reality-Tour für den Botanischen Garten Rostock. *AGIT: Journal für Angewandte Geoinformatik*.

White, S. and Feiner, S. (2009). SiteLens: Situated visualization techniques for urban site visits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1117–1120.

Wickham, H. (2010). A layered grammar of graphics. *Journal of Computational and Graphical Statistics*, 19(1):3–28.

Wilkinson, L. (2005). *The Grammar of Graphics*. Springer New York.

Willett, W., Aseniero, B. A., Carpendale, S., Dragicevic, P., Jansen, Y., Oehlberg, L., and Isenberg, P. (2021). Perception! immersion! empowerment! superpowers as inspiration for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):22–32.

Xie, B. (2013). *Head-related transfer function and virtual auditory display*. J. Ross Publishing.

Yang, S. and Sheedy, J. E. (2011). Effects of vergence and accommodative responses on viewer's comfort in viewing 3d stimuli. In *Stereoscopic Displays and Applications XXII*, volume 7863, pages 231–243. SPIE.

Yang, Y., Cordeil, M., Beyer, J., Dwyer, T., Marriott, K., and Pfister, H. (2020a). Embodied navigation in immersive abstract data visualization: Is overview+ detail or zooming better for 3d scatterplots? *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1214–1224.

Yang, Y., Dwyer, T., Marriott, K., Jenny, B., and Goodwin, S. (2020b). Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519.

Yu, D., Tang, L., Ye, F., and Chen, C. (2021). A virtual geographic environment for dynamic simulation and analysis of tailings dam failure. *International Journal of Digital Earth*, 14(9):1194–1212.

Yuan, M. (1999). Representing geographic information to enhance GIS support for complex spatiotemporal queries. *Transactions in GIS*, 3(2):137–160.

Zaker, R. and Coloma, E. (2018). Virtual reality-integrated workflow in bim-enabled projects collaboration and design review: A case study. *Visualization in Engineering*, 6(1):1–15.

Zeng, L., Weber, G., et al. (2011). Accessible maps for the visually impaired. In *Proceedings of IFIP INTERACT 2011 Workshop on ADDW, CEUR*, volume 792, pages 54–60.

Zenner, A. and Krüger, A. (2017). Shifty: A weight-shifting dynamic passive haptic proxy to enhance object perception in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1285–1294.

Zhang, C., Zeng, W., and Liu, L. (2021). UrbanVR: An immersive analytics system for context-aware urban design. *Computers & Graphics*, 99:128–138.

Zhang, L., Zhang, L., and Xu, X. (2016). Occlusion-free visualization of important geographic features in 3d urban environments. *ISPRS International Journal of Geo-Information*, 5(8):138.

Zhao, J., Simpson, M., Wallgrün, J. O., Sajjadi, P., and Klippel, A. (2020). Exploring the effects of geographic scale on spatial learning. *Cognitive Research: Principles and Implications*, 5(1):1–18.

Zollmann, S., Hoppe, C., Kluckner, S., Poglitsch, C., Bischof, H., and Reitmayr, G. (2014). Augmented reality for construction site monitoring and documentation. *Proceedings of the IEEE*, 102(2):137–154.

List of Figures

List of Tables

Eidesstattliche Erklärung

Hiermit erkläre ich durch eigenhändige Unterschrift, die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben. Die aus den Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Dissertation ist in dieser Form noch keiner anderen Prüfungsbehörde vorgelegt worden.

Unterschrift:

Ort, Datum: