AN ABSTRACT OF THE THESIS OF

Aditya Gune for the degree of Master of Science in Computer Science presented on June 19, 2018.

Title: Enabling Virtual Globes for A Visual-Auditory User Experience of Spatial Data

Abstract approved: ______________________________

Raffaele De Amicis

Geographical datasets are large, complex, and can be difficult for users to navigate and derive meaning from. These datasets, as well as the unique insights derived from them, provide tremendous opportunity for social change – many of the global challenges humankind is currently facing can benefit from analytics or visualization tools applied to geospatial data. Previous research has introduced platforms for storing and visualizing geospatial datasets, but lacks a platform for presenting this information in an immersive and multimodal way that enhances the user's cognition of the data. In this thesis, I investigate the challenge of, and propose a solution for, displaying and interacting with geospatial data through visualization and sonification. This solution enables users to create a multi-modal representation that is capable of empowering interactive geospatial data exploration. The methodology is integrated into a fully interoperable, standards-compliant spatial data infrastructure, ensuring that the method can be applied to a variety of datasets and application models to tackle a wide range of real-world challenges.

Enabling Virtual Globes for A Visual-Auditory User Experience of Spatial Data

by

Aditya Gune

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 19, 2018
Commencement June 2019

Master of Science thesis of Aditya Gune presented on June 19, 2018.

APPROVED:

Major Professor, representing Computer Science

Director of the School of Electrical Engineering and Computer Science

---

Dean of the Graduate School

Aditya Gune, Author

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor, Dr. Raffaele De Amicis, for his patience and guidance. His advice and support has been invaluable in the development of this thesis, as well as my personal and professional growth. I am also very grateful to my graduate committee: Dr. Margaret Burnett, Dr. Christopher Sanchez, and Dr. Kyle Niemeyer for their interest in my work.

I owe tremendous thanks to my family: my parents, Nitin and Vibha, my aunt Ranjana, my uncle Sandeep, my cousin Sourabh, for their love, understanding, and support. Your example inspires me to reach for higher successes.

Finally I would like to thank my close friends Vee, Sriram, Hannah, and Sara, for always being there for me through this amazing adventure.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Chapter 1: Introduction

Geospatial data is growing at a tremendous rate – sensor enriched buildings, streets, and city features create large amounts of geodata every day [2]. Geodata is by nature large and complex, and many of these datasets are impossible to understand without extensive filtering and sorting, which places a heavy cognitive load on the user. The concept of spatial data infrastructures (SDIs) evolved to address this issue, allowing spatial data to be stored in standardized databases and schemas that could serve as the basis for visual displays of data [3]. Modern SDIs support web-based visualizations of geospatial data, often using 2D map layers as a visual display. Some presentations also make use of more interactive virtual globe technology [4]. However, the large, complex, and varying nature of geospatial datasets shows that such visualization techniques are insufficient to properly explore spatial data [5, 6]. Furthermore, the increasing application of techniques such as machine learning and artificial intelligence generates more penetrating insights based on large geospatial datasets. Machine learning techniques have been applied to geospatial data in order to predict flood risks, epidemic outbreaks, forest health, climate change, and more [7, 8, 9, 10]. At the same time, the ability to present these insights to users in a meaningful way remains a challenge. Research such as Oregon State University's Explainable AI initiative attempt to provide greater transparency in how these algorithms work; however, presentation of complex geographical datasets, as well as the insights derived from them, must serve the user's needs by enhancing their cognition of the information [11].

These large spatial datasets, as well as the unique insights derived from them, provide tremendous opportunity for social change – of the 15 global challenges humankind is facing, as many as half can benefit from analytics or visualization tools applied to geospatial data [2]. The need for user-centric displays of geographical information is underscored by the opportunity the data presents. Presentations that augment the user's cognition of geographical data can support decision-makers in addressing critical challenges such as hazardous waste management and more [12]. These challenges demand a new generation of tools that empower decision makers by providing more immersive displays of geospatial data that enhance the user's understanding.

Failing to meet this need could have serious consequences, including potentially lowering public trust in institutions such as government and academia. Marzouki et. al. noted that citizens with low levels of participation in government showed higher distrust, and that increasing transparency through geovisualization tools had increased citizen involvement [13]. Similar phenomena have been observed with regards to community involvement in decision making with regards to urban features such as sidewalks and road design [13]. Clearly, failing to meet the need for a transparent, user-centric platform for geovisualization – one that enhances user cognition of the geographical data – has negative impacts throughout government and society.

Our contribution is a step forward in meeting this need, and supporting better decision making by enhancing user understanding of complex spatial data. Further research efforts should provide tools that encourage greater accessibility to spatial data and that improve the cognition of decision makers who use this data. This twofold goal – making complex spatial data more accessible and more understandable – is fundamental to addressing the challenges noted above [14].

Our work introduces a methodology and spatial data infrastructure that enables users to experience large geospatial datasets in a multimodal way. Our method applies sonification to the semantic data associated with geospatial datasets, such as population, building height, and other properties associated with urban features. This supports enhancement of the user's cognitive understanding of not just spatial data, but the real-world information contextual to the spatial data.

Previous work has demonstrated the importance of sonification for improving the user's understanding of data, especially of temporal patterns within a dataset [15, 16]. However, there has been little work on creating immersive multimodal displays using sonification for the purpose of enhancing user cognition of spatial datasets [16]. Our work fills this gap by introducing a system architecture and methodology that supports a multimodal presentation of spatial data for cognitive enhancement.

Our spatial data infrastructure uses standards-compliant spatial databases and schemas to store and process geospatial data in an extensible, interoperable way. We collect large geospatial datasets into spatial databases, and specify their subdivision into a series of smaller subdatasets, or "tiles", which are then converted to a web-compliant format for 3D visualization. This allows us to restrict the number of spatial data features encoded per tile, and control which features and data points are visualized and sonified. This control is crucial to the effective enhancement of the user's understanding, because failing to control the subdivision for spatial data would present extraneous information, and confuse or mislead the user. Our contribution also introduces a methodology for this multimodal approach, allowing users to visualize their own datasets and apply custom sonification algorithms to create their own multimodal representations.

Chapter 2: Problem Statement

The purpose of this thesis is to present a methodology and spatial data infrastructure capable of supporting the effective display of geographical information through a visual and auditory medium. Our central hypothesis is that such a multimodal representation can enhance the user's cognition of spatial data. The integrity of this hypothesis is supported by research validating that a multimodal presentation of sonification and visualization can enhance the user's understanding of data [5, 6, 17].

To accomplish this purpose, our research has three specific aims:

1. The first aim is to preserve the coherence between the user's visual and auditory perception of the data, meaning the data that the visualization and sonification display should be congruent and change in response to user input.

2. Secondly, the methodology should include an interaction method that allows users to manipulate the parameters of the sonification interactively at the runtime level. The user should also be able to consume data at a self-determined pace, rather than having all the data presented to them at once. This user-centric delivery of data is based on the *segmentation principle*, introduced by research in cognitive learning theory. The segmentation principle states that users are better able to learn when they can control the flow of data delivered to them, so that their cognitive processing of information is not interrupted by an uncontrolled flow of new information that they must perceive [18].

3. Finally, our approach should meet requirements necessary for addressing the broader challenges that geospatial data can be applied to. This means that spatial data must be stored and processed through interoperable federated data infrastructures, and added-value processes such as sonification must be delivered through a service-based software component.

Fulfilling these aims will allow us to enable a presentation of geographical data that can empower exploration and enhance the user's cognition of the dataset.

In creating a solution that fulfills these goals, the first question we address is the extent of geographical data that should be presented to the user. Our approach is that creating a display whose sonification is congruent with the information under the user's visual attention can augment the user's cognition of the dataset. The sonification algorithm we use therefore sonifies the data within the user's field of view.

This approach for determining the extent of data to be sonified is based on prior research in cognitive theory known as the *coherence principle*: users are better able to learn when extraneous information has been excluded from their view [18]. Applying the coherence principle to a multimodal display means that neither the sonification nor the visualization should introduce extraneous data. A sonification that is not congruent with the visual focus of the user would risk presenting unrelated information, seriously impacting user cognition of the dataset.

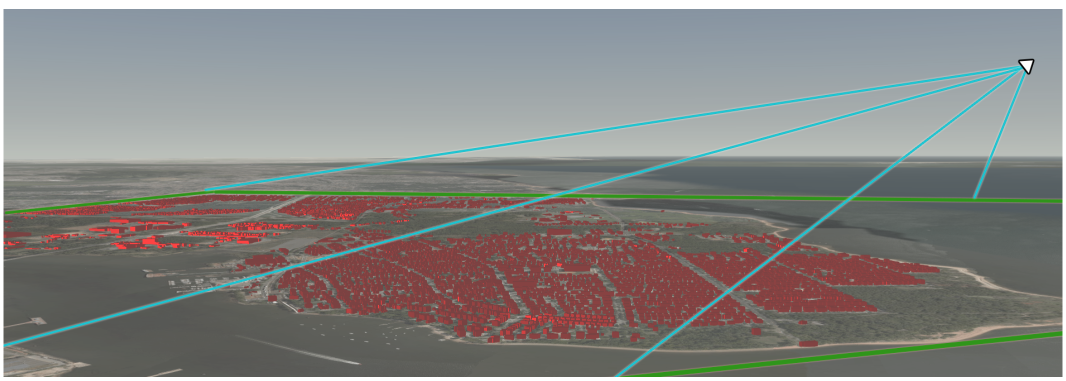

Figure 2.1: Everything within the user's field of view (blue) must be sonified to ensure that the user's visual and auditory perception of the data are congruent.

Because sonification is temporal in nature, we must also consider which order this data should be presented in. One possible option is to sonify based on proximity to the user. That is, we can say that the focus of the user's attention is the closest spatial point visible to the user. Alternatively, we can choose to sonify spatial features based on specific characteristics of the data that catch the user's attention, such as sonifying the tallest skyscrapers first.

In either case, we restrict the amount of data being presented to the user, and sonify and visualize only relevant data that corresponds to the geographical area within the user's field of view. Cognitively, this allows us to draw the user's attention to the data

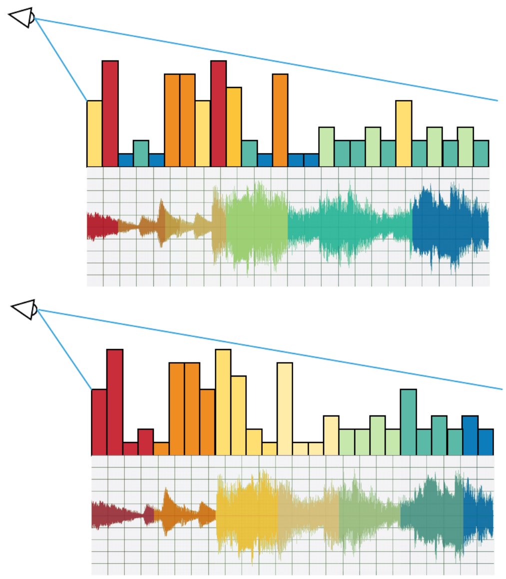

Figure 2.2: Schematic of sonification by height (top) and proximity (bottom). Sonification by proximity clusters buildings by spatial distance, whereas sonification by height plays the tallest buildings first, regardless of how close or far they are to the user.

by eliminating extraneous information which might distract them from the data [18].

Our methodology also meets the requirements for interoperable solutions for broader needs. We integrate this approach into an interoperable spatial data infrastructure consisting of a data layer for storage and processing of spatial data, a middleware web service for sonification, and a presentation layer application that can enhance the user's understanding.

The SDI's data layer uses CityGML schemas for storage and processing of spatial datasets. CityGML is an interoperable XML-type standard specified by the Open Geospatial Consortium for storing and transferring urban datasets [19]. Though CityGML is not optimized for streaming and visualization, previous research has successfully visualized CityGML spatial data using other transmission formats such as glTF [1, 20].

The presentation layer contains the Cesium 3D virtual globe for providing a visual display and the main point of user interaction with our application, which presents urban data in Cesium using a glTF-based format for visualizing large 3D spatial datasets [21]. The glTF specification is an interoperable standard for transfer and runtime delivery of

3D content [20, 21, 22].

Our SDI also provides sonification through a web service in the middleware layer. The web service takes HTTP requests containing semantic data, and outputs a sonification that is played in real time by the presentation layer to create a coherent multimodal representation that can enhance the user's cognitive understanding of the data.

The structure of spatial data infrastructure, including the standards used, will be covered in more detail in chapter 3.

Chapter 3: Spatial Data Infrastructures

Spatial data infrastructures support the presentation of geographical information by providing a set of standards and frameworks that allow for storage, processing, visualization, and sharing of geospatial data [23]. Current spatial data infrastructures support both spatial (map-based) and semantic (text-based) searching by the user, allowing users greater freedom of exploration. The rapid democratization of data and rise of non-expert volunteered geographic information has increased the importance of SDIs in providing greater accessibility and user understanding of geospatial datasets [24].

3.1 How SDIs Evolved

Due to the inherent complexity of geospatial data, presenting large geographical datasets in a way that users can explore and understand is a challenge [25]. Historically, this information took the form of paper maps, which computer science and information technology replaced with geographic information systems (GIS) – software systems which stored and presented geographical information with the purpose of analyzing spatial information and modeling spatial processes [23].

Many early GIS were developed as either intraorganizational purpose-built solutions designed to make existing tasks more easier, or as enterprise solutions intended to facilitate customer satisfaction [3]. However, the increasing amount of geospatial data being generated created a demand for a standardization of the technology used to store and present this data. The concept of spatial data infrastructures was introduced in the 1980s by national survey and mapping agencies, in an effort to standardize geospatial information and reconcile differences between existing technology standards and schemas [3]. By 1994, the US Government's Federal Geographic Data Commission had been established to set up a National Spatial Data Infrastructure. As part of initiative, the Open Geospatial Consortium (OGC) was founded to establish standards for geospatial applications. Other countries also developed their own efforts, such as the Japanese NSDI in 1999, the Indian NSDI in 2005, and the European INSPIRE initiative in 2007

[26, 27, 28]. Each of these SDI initiatives developed around a specific goal. The US NSDI developed to support collection of large amounts of spatial data [28]. In Japan's case, the NSDI evolved out of a need for better coordination during hazards planning – a response to the 1995 earthquake in Kobe, Japan [27]. The Indian SDI evolved with the goal of supporting increased citizen access to spatial data [26].

In all cases however, a principal factor in meeting these goals was the need to standardize the collection, storage, and availability of geospatial data in support of the organization's efforts to address broader national and international challenges.

This drive for standardization emerged from the need to store and process geospatial information in an interoperable way. That is, geospatial data should be stored and processed in such a way that it can be accessed and used across a wide range of systems [23]. This requirement for interoperability is driven by the diversity of data models and sources from which spatial data is taken [29]. Diversity of data often leads to a hetereogeneity of formats in which geospatial data is stored, as many GIS systems use proprietary (and sometimes purpose-built) software, data models, or databases for storing geographical information [30]. These differences in the way data is stored lead to incompatibilities between GIS systems, and force users to perform complicated data transformations. Often, these conversions require users to either purchase or develop tools, costing time and money due to a lack of interoperability in data sharing. The acquisition of tools for usage from heterogeneous data sources is a significant resource investments for users [30]. This resource demand can reduce accessibility of geospatial data and can impede users' effective use of GIS systems.

3.2 Architecture of an SDI

SDIs can have a system architecture consisting of three tiers, or layers, of software components for storing, transferring, and presenting spatial data. The information or data tier is responsible for storing the spatial data to be viewed. In many cases, the data tier contains spatial databases to store and process spatial data. The middleware tier contains web services or other software components that provide value-added services to the user, and the presentation tier contains application logic and code that allows the user to visualize and interact with geospatial data. The presentation layer often takes the form of geoportals, which provide web-based interfaces for searching and exploring geospatial information [24].

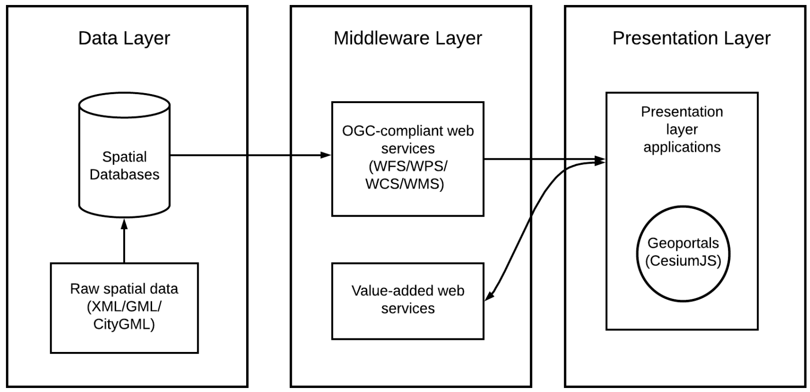

Figure 3.1: A three-tiered spatial data infrastructure consists of a data layer, a middleware layer, and a presentation layer. Diagram adapted from Prandi et al [1].

3.2.1 Data Layer

The data tier stores and processes geographic information through the use of spatial databases, which use standard database models like the relational database whose schemas have been augmented to store and modify spatial data [31]. Spatial databases allow the database to store spatial and geometric data such as points, lines, and surfaces. This support for spatial data types is critical, as it ensures that spatial data is stored accurately and without loss of detail [32]. Spatial databases also provide functionality for organizing and modifying data through the use of spatial queries and functions [31]. This ability to define processes for handling and manipulating data – such as converting between coordinate reference systems – provides greater accessibility and accuracy by allowing users to overcome issues of data diversity [32]. The data tier can also store spatial and semantic elements – such as 3D model geometry – at the file level [1].

3.2.2 Middleware Layer

SDIs also contain a middleware layer, which acts as a bridge between the information and application layers. This layer often hosts web services or other software components which allow applications in the presentation layer to access spatial information in the data layer, either from spatial databases, or from file-level storage [1]. Middleware-level web services can also be used to present value-added services. For instance, with the rise of technologies like parallel computing, high performance computing, and advanced data analytics, the middleware tier can feed data to third-party computing services to generate real-time insights for users based on geospatial data [28].

3.2.3 Presentation Layer

The presentation layer is the main point of user interaction. This tier contains all application code, including geographical interfaces like virtual globes. The presentation layer receives user input, retrieves data from the data layer through the middleware layer, and presents data to the user in an interactive and effective way [1].

Virtual globes in the presentation layer are often the main point of user interaction with the SDI. One of the key benefits of virtual globes is their ability to display multiple types of geographical data and combine various data streams to provide the user with a more realistic representation of real-world features [33]. The Google Earth virtual globe allows users to visualize their own geospatial data using the Keyhole Markup Language (KML) standard. Similarly, the open-source Cesium virtual supports multiple data formats, allowing users to explore their own data. Cesium offers a JavaScript interface that allows developers to create terrain and feature rich web applications to visualize geodata. For instance, it supports layer imagery from multiple sources, as well as geometries such as points, polylines, and polygons. Its open-source nature means users can extend its functionality for their own applications and use cases, promoting accessibility and user exploration of geographical datasets.

In addition to virtual globes, the presentation layer can also support other applications for displaying data to the user, such as 2D geoportals and web graphical interfaces. These applications can use community accepted APIs such as the Google Maps API, or interoperable standards such as OpenLS, for displaying geospatial data to the user. This layer can also be device agnostic, displaying data on mobile devices and desktops alike [34]. The ability to present data in multiple ways provides greater accessibility of spatial information for users.

3.3 Standards

Interoperability and data integration in SDIs can be achieved in a variety of ways. One of the most common is the application of pre-defined and community-accepted standards and practices to resolve inconsistencies in spatial data. These standards and practices are usually introduced by national or international organizations such as the Open Geospatial Consortium or the International Standards Organization. For instance, many GI systems use GML to store spatial data [23]. GML is a version of the XML schema that can encode spatial data, such as point, line, and polygon geometry. GML models real-world geographic features as a collection of geometries following a hierarchical data structure that is searchable by feature. GML also specifies the SRS, or spatial reference system (also known as CRS, or coordinate reference system) for spatial datasets. A CRS is a set of coordinate system axes that relates to the Earth through a datum defining the size and shape of the Earth. This relationship between axes specifies how coordinates in one coordinate system can be transformed into coordinates in another system [19]. Because of its widespread adoption and certification by the OGC, the GML standard allows SDIs to access and combine spatial data from a variety of sources, greatly increasing interoperability [35].

Other interoperable standards include CityGML – a version of GML designed to store 3D city models. CityGML was introduced in 2005 due to a lack of structural interoperability in 3D city models [36]. That is, languages and schema such as GML allowed sharing of data between heterogeneous forms of storing spatial data (such as spatial databases); however, geographical features could be labeled differently based on the discipline and application [37]. CityGML established this structural interoperability by introducing a set of common syntactic labels for objects in 3D city models [36], allowing for the sharing of spatial data for urban environments without the risk of inconsistent labeling. The CityGML schema is capable of encoding features commonly found in urban environments, such as trees, bridges, roads, and buildings [19].

However, while CityGML is capable of storing geospatial data, it is inefficient for visualizing it [1]. This is in part because it inherits certain concepts from GML, such as its tendency to use complex polygons for modeling surfaces. To easily present and transmit 3D data, the Khronos Group – a non-profit responsible for standards in 3D computer graphics – released glTF (GL Transmission Format), an interoperable format for presenting 3D content through the JSON standard. glTF files encode camera and scene information, the geometry of the 3D object, textures, animations, and more. Researchers have used a combination of CityGML and glTF to stream large urban datasets in real time [20].

In addition to data models and formats, SDIs can use standardized services to handle user requests and serve data. The OGC's Web Feature Service standard is an XML based standard that queries spatial databases of any format and returns spatial data in GML format. The use of GML format allows users to associate data from multiple WFS requests to different databases. WFS handles HTTP requests and allows create, read, update, and delete (CRUD) operations, including querying data based on spatial and non-spatial parameters. Critically, WFS allows querying of spatial databases at the *feature* level, avoiding the costs associated with data exchange at the file level [35].

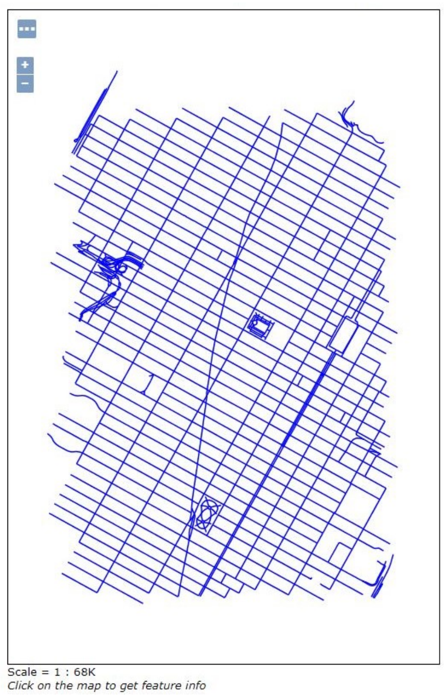

Figure 3.2: A WMS response of roads in New York City.

Other OGC-compliant standards include Web Processing Service (WPS) and Web Map Service (WMS). WPS standardizes how geospatial processing services should handle requests and responses. WPS allows the user to retrieve or transform data in real time through a client, allowing users to generate new information from existing datasets [38]. WMS standardizes the creation and display of maps and geospatial imagery. It uses standard image formats such as SVG, PNG, GIF, or JPEG, and allows users to specify parameters such as spatial reference system and size.

The adoption of these community accepted standards provides for greater interoperability by defining common schemas and methods of storage and access, and ensures that diversity of data will not impede user access, use, and exploration of geographical information.

Chapter 4: Architecture and Methodology for a Multimodal Display

Our SDI follows the architecture described in the previous chapter, with a data layer for storage, a middleware layer for services, and a presentation layer for interaction.

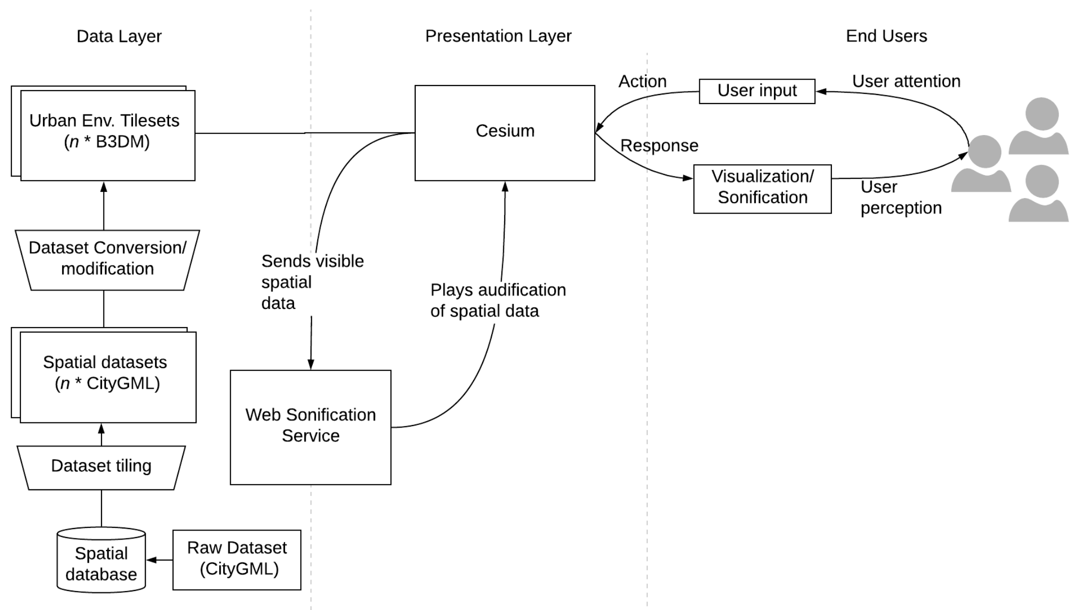

Figure 4.1: System architecture for a multi-modal representation of an urban environment.

The data layer consumes spatial data in CityGML format and stores it in a PostGIS spatial database system. This spatial database is further extended using 3DCityDB's database schema tool, which augments the database schema by adding spatial data types and support for data transformation, import and export options [39] for CityGML spatial datasets. Using CityGML and 3DCityDB allows us to efficiently store spatial datasets and to transfer spatial data in an interoperable way. Furthermore, it allows us to retrieve, manipulate, and export data along spatial dimensions (such as exporting one geographical area of a city). The data is then converted to the B3DM format and passed to the presentation layer for visualization.

The presentation layer visualizes spatial data through the Batched 3D Model (B3DM) format, a specification introduced by the Cesium team for visualizing large 3D spatial datasets. B3DM encodes both the glTF-based 3D content, as well as any semantic data associated with it (such as building height and address) [21]. Our application identifies the geographical areas under the user's visual attention and passes the corresponding semantic data encoded in the B3DM to a middleware level web service for sonification. This sound data is returned to the presentation tier application for a combined visual and auditory experience.

The presentation layer also allows users to interact and select which data to experience by navigating the virtual globe. However, to ensure our platform can support a multimodal experience capable of enhancing the user's cognition of the dataset, we must meet the goals described in our problem statement: ensuring coherence between the visual and auditory components, and presentation of data at a segmented, user-driven pace. To accomplish these goals, we introduce a methodology to manipulate and process data in a way that can meet these requirements.

4.1 Methodology

We process the spatial dataset for this presentation by converting it from CityGML to the glTF-based Batched 3D Model (B3DM) format. B3DM specification encodes a dataset as a hierarchical collection of B3DM tiles, with each tile containing a number of spatial features, such as buildings, roads, or other objects. Each tile contains a feature table, which carries the geographical data of every building within the tile; a batch table, which holds semantic data such as building area; and the glTF 3D geometry for the tile's features [21]. Our application "captures" which tiles are visible to the user, and reads the semantic data from those tiles' batch table; this data is passed to the web service for sonification. Tiles are arranged in a tree data structure, with child tiles always falling within the geographical area of their parent tile.

In large datasets like NYC, B3DM tiles are often deeply nested, and child tiles are not loaded at the same time as the parent. This results in a data display where much of the data is "pre-filtered" out from the user's view – that is, the user could view a geographical area, but only be able to perceive some of the spatial data, providing them an incomplete display. For instance, in Figure 4.1, much of the data on the right half of the user's view has been filtered out and is not visible.

Figure 4.2: An example of how the user perceives data that has been pre-filtered based on the B3DM hierarchy.

The non-visible data is encoded in tiles that are not being presented to the user due to the tiles' position in the B3DM hierarchy. As a result, the user's perception of the data does not match the area that has their attention.

The user's natural response to only perceiving some data would be to navigate closer. However, this presents another problem – the user would be viewing both the child tile and its parent tile, which is geographically much larger. In such a case, the sonification service would sonify data associated with all visible tiles, presenting the user with large amounts of data that was outside the field of view. Effectively, the user would be perceiving extraneous information aurally, resulting in a lack of congruence between the user's visual and auditory perception. For example, in Figure 4.2, the user only perceives features in tile B when they navigate close to it. However, because B's parent is also visible, the implementation sonifies all features in both tiles B and A, many of which are outside the user's field of view. As a result, the user either perceives insufficient amounts of data, or too much of it.

We address this issue by subdividing the initial dataset to create smaller, non-hierarchical B3DM tilesets for more accurate visualization and sonification. We specify the tiling schema of our raw CityGML data in the information layer to control the num-

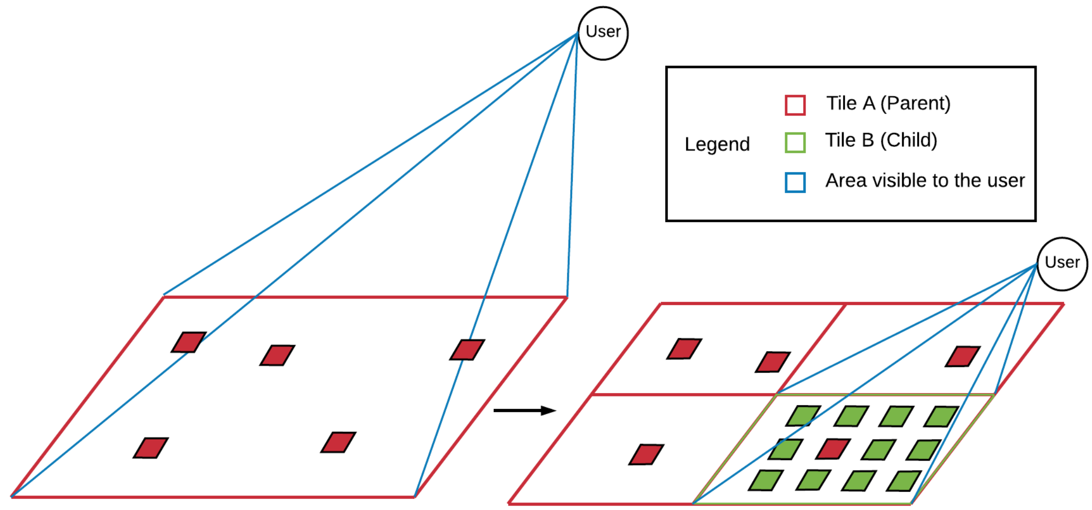

Figure 4.3: The user is not able to perceive the features in Tile B until they navigate closer, at which point the system sonifies both Tiles A and B.

ber of features encoded in each tile, and to ensure that there will be no spatial overlap between tiles. This is accomplished by first storing the entire CityGML raw dataset into a 3DCityDB spatial database, and using the 3DCityDB to export a series of uniformly sized CityGML subdatasets. This method allows us to control the extents and number of features for each tile, effectively controlling the size at the lowest level of data. This ensures that the final B3DM tiles output by our method will be no larger than those we specified in our CityGML subdivision. The small tile size ensures that the user's experience of the visualization will not include large amounts of extraneous data that adds to the user's cognitive load.

Each CityGML dataset is then converted into a B3DM tileset containing only one B3DM tile. After conversion to B3DM, we apply a viewer request volume value to each tileset. The viewer request volume is a property of B3DM tiles that defines the volume of 3D space in which the camera must be for the tile to be rendered. Setting this property allows us to ensure that only those features visible and close to the user's field of view are rendered, eliminating extraneous data from being presented to the user.

Our tiling method effectively restricts the number of features a CityGML file or B3DM tileset can contain, and restricts each tileset to one tile in place of the usual hierarchy of tiles. By controlling the size of the visualized area, this method of creating a single-level hierarchy of small tiles further supports our goal of introducing a platform that supports the enhancement of user understanding. It eliminates the possibility of both the pre-filtering of data from the user's view, as well as the tile overlap, where the user is presented with visual and aural information from multiple tiles. This organization of spatial information allows users to fully perceive, explore, and enhance their understanding of the data that has their attention.

4.2 Sonification

Sonification has been shown to effectively convey both spatial and non-spatial data to users [15, 16, 40]. Presenting such data as sound can give users a new perspective, leading them to unique insights. For instance, Hogan and Hornecker found that people appear to rely more on intuition and real-world experiences when making sense of haptic and auditory representations, compared to visual representations, which are experienced more as pragmatic tools [41].

We use sonification to present semantic information associated with spatial data. Aside from geographic location, urban datasets often encode large amounts of non-spatial data. An urban dataset could model building attributes such as height and area, as well as temporal data such as traffic conditions or crime patterns. We present this data by creating an auditory graph, which serves as an aural version of a visual plot or line graph [42]. Our system allows the application of a mathematical function ϕ to map data points to musical notes. For a given visualization, we consider a source matrix S consisting of all data associated with the urban environment. The choice of mapping function, and the range of data points which are sonified, are both driven by user input and attention.

Figure 4.4 presents a sample mapping function: a linear transformation, where each value within the input domain is linearly mapped to a value in the output domain.

$$\phi(X,Y) = (x_i - min(X)) * \left(\frac{range(Y)}{range(X)}\right) + min(Y)$$

$$-\infty \leq X \leq \infty$$

$$36 \leq Y \leq 86$$

Figure 4.4: This figure presents a linear transformation mapping function, where each value within the input domain is linearly mapped to a value in the output domain.

Here, X represents a subdomain of data values chosen from the full range of data points (such building area values ranging from 50 to 300), and Y represents an output subdomain within the range of notes, where both subdomains are numerically specified by user input. Figure 4.5 provides a schematic of how the user can define sonification properties.

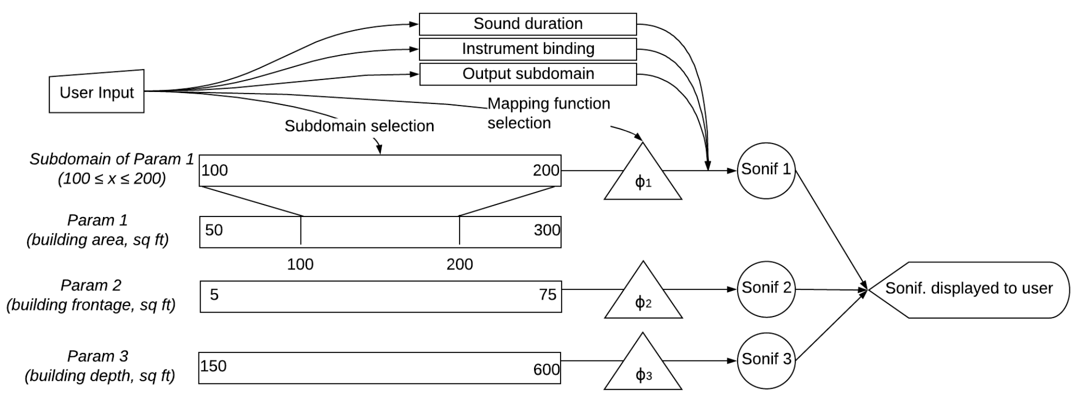

Figure 4.5: The user can provide inputs through a control panel to select a subdomain of data, a mapping function, and sound parameters.

Allowing the user to choose the mapping function and range of data to sonify ensures our sonification method is user-centric. It allows the user to understand which data they perceive as sound and define the relationship between the data and the sound they perceive. This provides them the ability to make connections between the visually and aurally presented data, which can enhance their cognition of the data.

The sound created by the sonification is passed back to the application and is played for the user in synchronization with a visual color change. This creates coherence between the visual and auditory elements of the display and draws the user's attention to the geographical area being sonified. It further allows users to make associations between the data they perceive visually and the data they perceive aurally, allowing users to increase their cognition and understanding of the dataset.

Chapter 5: User Interaction

In developing our methodology, we also implemented an interaction schema – a concept version of a user interface capable of supporting the type of multimodal display our methodology enables. The browser-based user interface is composed of two main elements: the Cesium container, where users can navigate the digital globe, and the menu-style control panel, which allows users to tune the parameters of the visualization and sonification. Within the Cesium container, users can navigate the digital globe and interact with the visualization component of our representation, including picking features to gain further information. The system presents spatial data both aurally and visually when users navigate the globe to visualize a geographical area. This user-driven trigger for displaying data ensures that the user is not overwhelmed by an uncontrolled flow of new data, and can choose when to perceive new information.

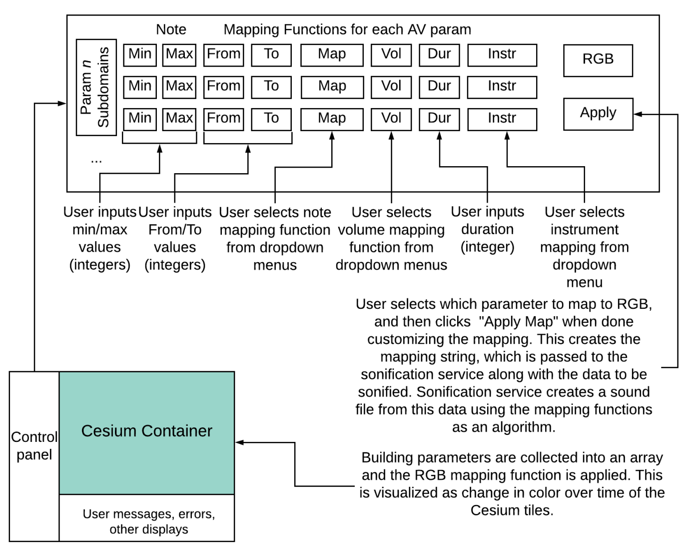

Figure 5.1: Our user interface allows users to manipulate data and sound parameters for fine control of both the visualization and sonification.

The control panel contains the input fields allowing users to specify which data parameters to sonify, and how they will be mapped to sound parameters. For any data parameter, users can choose a subdomain within the range of values for that parameter and map them to a range of numbers representing musical notes. For a cluster of buildings that has building area values ranging from 50 to 300, the user would be able to specify subdomains such as $100 \leq sub_i \leq 200$. The data from each subdomain is then mapped to a note range (which the user can also specify) using a mapping function, which the user chooses from a drop-down menu. A user can choose to map the data values from $100 \leq sub_i \leq 200$ to any subdomain within the range [36, 86], which represents the range of audible notes when converted to MIDI numbers. The resulting MIDI number is then played by the sonification service using a MIDI player. This ability to perceive atomic elements of data, such as a single data parameter, enables users to explore the effects of individual parameters or data points through sonification.

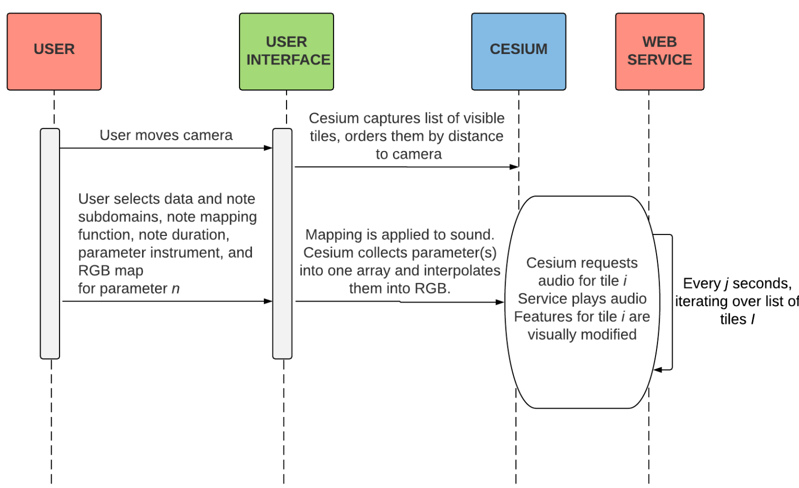

Figure 5.2: A sequence diagram of how the user can apply our interaction schema to control the visual and auditory elements.

The ability to select sound parameters and a mapping function empowers the user to define the relationship between data and sound. Understanding this relationship is critical to effective use of the sonification for exploration and enhancement of the user's cognition. Furthermore, it allows the user to perceive the sound data in chunks, and at the pace they desire – a concept known in cognitive research as the segmentation principle. Our interaction schema allows the user to segment the information, so their perception enhances their learning and cognition of the dataset [18].

Chapter 6: Development Process

We developed our architecture and methodology incrementally, starting with an implementation capable of visualizing a large spatial dataset and sonifying the dataset's semantic information. We observed that many of the tiles seemed to overlap spatially, causing spatial data to be filtered out of the user's view.

To better understand how tile boundaries were formed and how tiles were hierarchically organized, we tested multiple visualizations. The first was a "color by tile" visualization which applied random colors to each B3DM tile within the tileset. We found that a given tile could contain multiple child tiles, and that each child tile encoded an area within the bounding volume of the parent tile. However, in many cases, features within a child tile's borders would in fact belong to its parent tile.

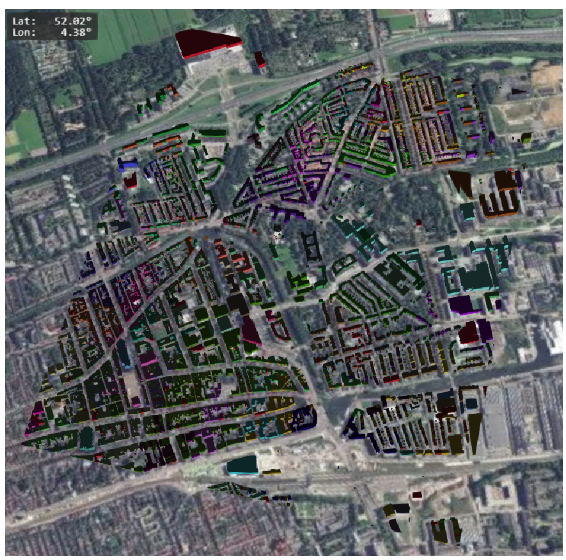

Figure 6.1: Our first visualization colored each building in the dataset based on which tile it belonged to, indicating the extents of each tile.

We found that larger datasets (such as NYC) had very deep hierarchies, with some tiles holding as few as 10 features, and others holding hundreds or thousands. This indicated that in order to support a representation capable of augmenting user cognition of the data, we needed to eliminate the overlap caused by the B3DM data structure.

Our second visualization was a "color by order" visualization, which colored B3DM tiles based on the order in which they were loaded. Our goal was to understand the order in which data was presented to the user. We used an RGB gradient to represent order: the first tiles loaded were red, later tiles were green, and the last tiles loaded were blue.

Figure 6.2: Our second visualization tested the order in which tiles were loaded.

We found that though the first tiles visualized were often close to the camera, deeply nested hierarchies would reduce the consistency of the load order. Often, a large amount of data would be filtered out of view of the user. This data would be inaccessible unless the user navigated closer. In order to support presentations that could increase the user's understanding of the data, we needed to be able to control the order in which the user perceived spatial information, and ensure that all relevant data within the user's field of view was presented.

Finally, we tested a "color by hierarchy" visualization, which applied an RGB spectrum to B3DM tiles based on their position in the tree data structure. For instance, the root tile, at the top of the tree, would be red; tiles exactly halfway down the tree would be green; and tiles that represented leaf nodes in the tree would be blue. Tiles in between these would be colored based on a linear interpolation that mapped depth to color values. We observed that the majority of features within the dataset were green, indicating a position in the middle of the hierarchy.

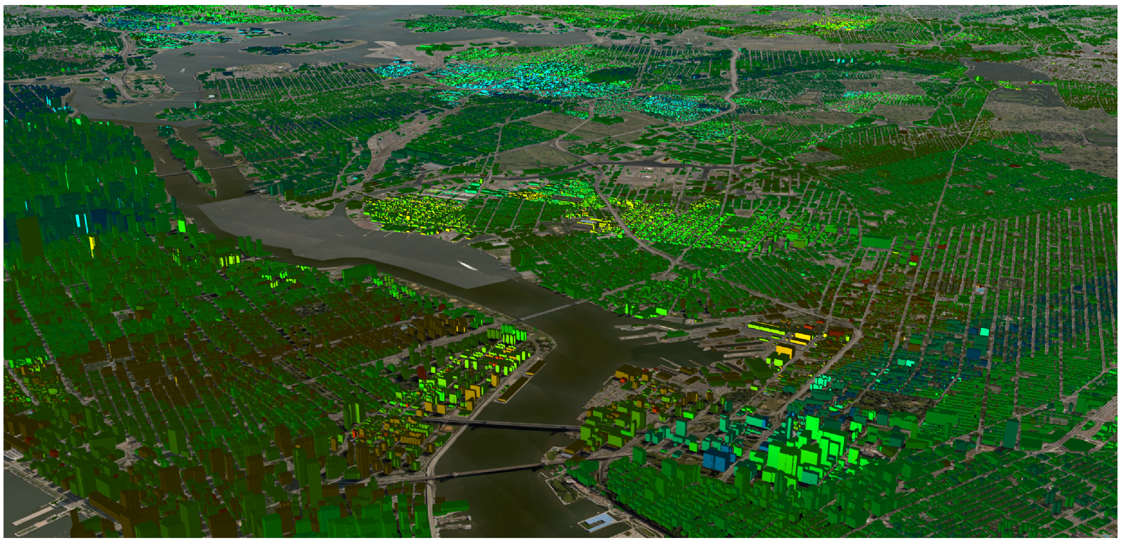

Figure 6.3: Our third visualization examined the position of tiles within the tree data structure.

Comparing this visualization with the color-by-order visualization, we observed that the tiles that were filtered out of the user's view were often deeper in the hierarchy. This indicated that the majority of data within a large dataset would be would be filtered out from the user's view, affecting their understanding of the information.

In all cases, the biggest obstacle was the nature of the B3DM tileset itself. Large tiles with deeply hierarchical structures presented significant problems for effective visualization and sonification. Often, this structure would lead to a large amount of extraneous data being loaded, such as areas of the dataset that were geographically distant from the location the camera was viewing.

Our solution to this problem is the method described above: the subdivision of the initial dataset to create smaller, less hierarchical B3DM tilesets for more accurate visualization. The small size and non-hierarchical organization of the tiles allows us to effectively render buildings close to the camera and capture the non-spatial data associated with them.

Chapter 7: Evaluation and Discussion

Our work introduces a methodology and SDI that enables virtual globes for a multimodal display of data capable of augmenting the user's cognition. We achieve this by addressing three primary research aims: the preservation of audio-visual coherence in our display; the user-driven nature of our display to ensure segmentation of data delivery; and the interoperability of our method through the use of an SDI.

7.1 Audio-visual Coherence

Audio-visual coherence is the congruence of data displayed through visual and auditory modalities. To effectively enhance the user's cognition of the data, these modalities should present information that is under the visual attention of the user, and should avoid introducing extraneous information unrelated to the user's attention [18].

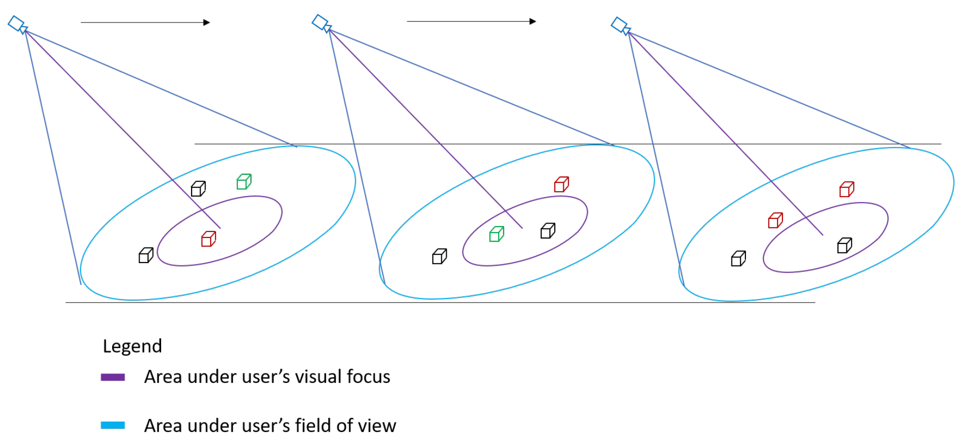

Figure 7.1: Most features are filtered out of the user's field of view.

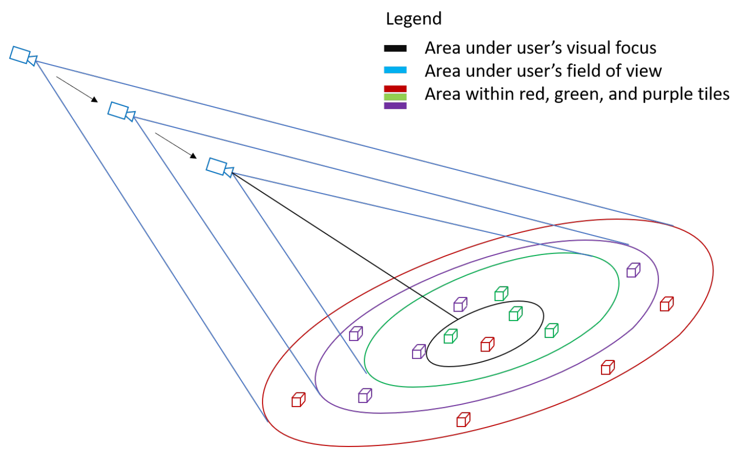

Figure 7.2: The user can only visualize higher-resolution features by navigating closer.

The default presentation of spatial data delivers information to the user based on inherent properties of the data, such as a feature's location in the data structure. However, this appears as to a pre-filtering of data, where the user's perception of data is not congruent with their visual attention. Furthermore, because the sonification is based on visible tiles, triggering the sonification would also fail to represent all data under the user's visual focus. A user who navigated closer to perceive data at a higher resolution would then potentially experience extraneous data due to the hierarchical nature of the tree data structure.

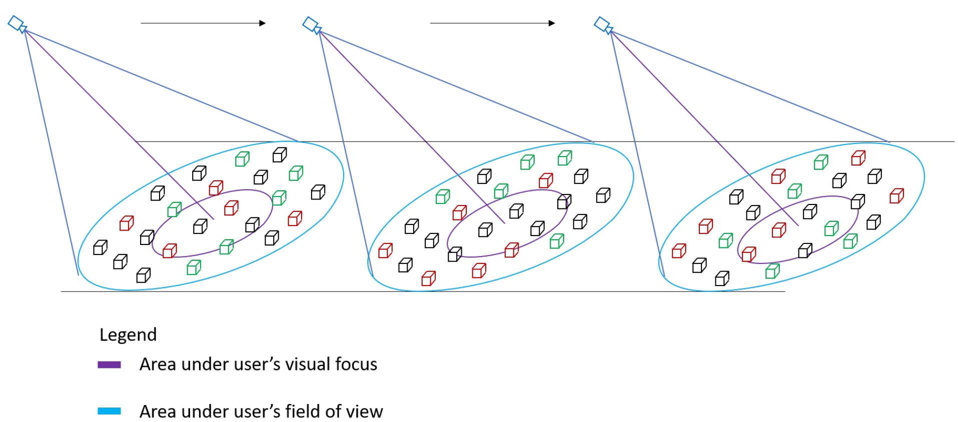

Figure 7.3: Our method displays *all* spatial features within a geographic area.

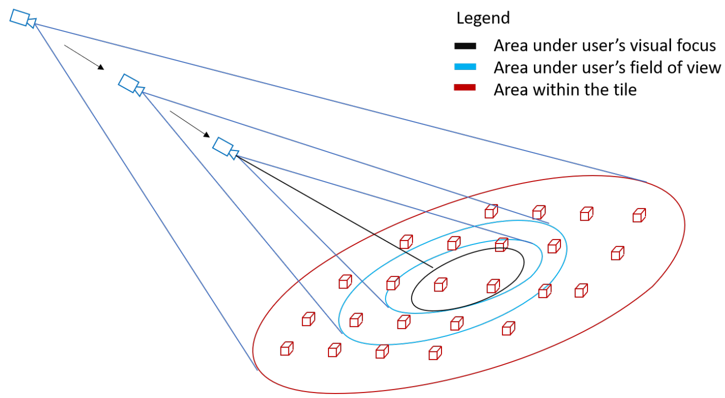

Figure 7.4: The user can navigate closer, but will still perceive all data in their field of view.

We introduce a method that enables a presentation of spatial data whose visual and auditory components are congruent with the user's visual attention.We manipulate the size and organization of spatial data by establishing a non-hierarchical data structure whose tiles are uniformly sized, eliminating the existing hierarchical structure with variable sized tiles. This uniform size and non-hierarchical nature eliminates the possibility of spatial overlap of tiles containing geographical features, and ensures that tiles contain data for all features within their geographical area.

This ensures that when the user navigates to a geographical area they perceive all tiles – and by extension all spatial data – within that area, guaranteeing that the user's perception of data matches their visual attention. Furthermore, our application of the view request volume property makes sure that only tiles within the user's visual perception. The view request volume defines the region of 3D space in which the user must be in order to be presented with data from a tile. By defining this property, we ensure that only tiles within the user's visual focus are rendered, avoiding the presentation of unrelated data.

The multimodal display captures only visible tiles for sonification, and this reorganization of data allows the system to render and sonify only tiles which have the user's visual focus, preventing extraneous data from being sonified. In this way, we control the spatial organization of data to ensure that the visual and auditory components of the display are congruent, and that the audio-visual presentation of data is coherent with the user's visual attention.

This audio-visual coherence is central to our goal of introducing a platform that enables presentations which can enhance user cognition of spatial data. Our method of enforcing audio-visual coherence ensures that presentations that use our platform will fulfill this requirement.

7.2 User-Driven Display

The ability of users to control when data is delivered is critical to enhancing cognition, as it makes sure that user understanding of current data will not be interrupted by delivery of new data. Allowing an uncontrolled flow of data to be delivered to the user could add to the user's perceptual load, and would violate the segmentation principle, which states that users learn better in self-paced chunks. For a multimodal presentation to enhance the user's cognition of data, it should allow the user to control when and how data is delivered [18].

To support presentations of data that allow segmented data delivery, we introduce an interaction schema and graphical user interface whose responses and delivery of data is driven by user input. Our user interface contains a virtual globe which can be navigated by end-users to visualize a geographical area. Presentation of data is triggered by this navigation, ensuring that new data will not be delivered to the user unless they actively provide input to perceive a geographical area.

Secondly, the interaction schema allows users to define parameters for sonification. The user can choose which semantic data parameters (such as building area or depth) should be represented as sound, and can select which values or subdomains within the data domain should be sonified. This allows users to segment the data delivery by choosing how much semantic data should be sonified, as well as to explore data in greater detail by accessing atomic elements of the data through sonification. For instance, users can choose to sonify one data parameter with one instrument to explore exactly how those values change across a dataset. The user can also determine the order in which data should be sonified. We introduce a method that displays data by proximity to the camera, but users can modify this presentation to prioritize properties of the data such as building area or importance.

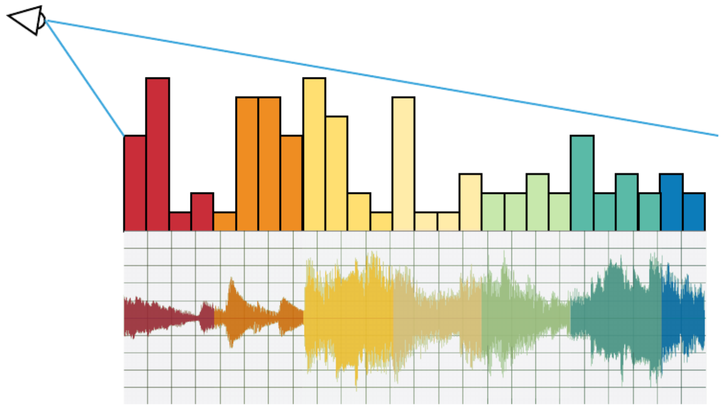

Figure 7.5: This multimodal display sonifies buildings in order of proximity to the user's camera. However, users can modify this to suit their own needs.

The interaction schema we introduce provides users with control over when data is delivered, through the navigational trigger; and how much of it is delivered, through the ability to choose data parameters and subdomains for sonification. By giving the end-user full control over the presentation of data, we ensure that our platform supports a display of spatial data that can enhance the user's cognition.

7.3 Interoperable Solutions

The ability to provide interoperable solutions is critical to the broader goal of enhancing user cognition in support of global challenges. Presentations of spatial data should promote accessibility and explainability of geospatial data, enhancing the cognition of users and decision makers. Effecting this requires providing an interoperable approach: platforms that enable these presentations should use federated spatial databases for storage and processing, and deploy value-added processes through web service-based software components.

We integrate the methodology for multimodal display into a spatial data infrastructure that uses interoperable federated spatial databases for storing and processing data. We also use community-accepted standards such as CityGML for storing spatial data at the file level, and use the glTF-based B3DM standard for presentation of spatial data as 3D content. Finally, the value-added process of sonification is delivered through a web service software component.

By fulfilling these three research aims, our contribution supports displays of spatial data that can enhance the user's cognition through visualization and sonification, and can provide a framework applicable to broader challenges surrounding geospatial data.

Chapter 8: Conclusion

The research goal of this thesis has been to explore how we can enable displays that augment the user's cognition of a geospatial dataset. Geospatial data is large, high-dimensional, and challenging to understand, even with powerful visualizations. The combination of visualization and sonification allows us to enable coherent, interactive presentations that can support a user's exploration and understanding of the data.

In developing this research we observed a number of limitations to the methodology. For instance, the performance of the system can decrease as the resolution of the B3DM tiles increases, as a function of the cost of loading data into the browser. This performance limitation has the potential to affect the user's perception and cognition of spatial data, and should be investigated further. A second limitation is the temporary de-synchronization between the auditory and visual modalities due to external factors such as network latency, hardware, or system performance.

Other findings include the need to extend spatial data infrastructures to enable presentations that enhance user cognition. We observed that accomplishing this requires thinking about spatial data infrastructures through the context of the user goals, rather than as a set of frameworks for storage and processing of data. SDIs should also serve user needs by providing the ability to fuse heterogeneous datasets to empower displays for enhancing user understanding of multiple data streams.

Additionally, the use of visual and auditory displays presents the need to investigate how best sonification can augment user cognition. Our interaction schema allows users to modify the display of spatial data based on their own cognitive requirements; however, pertinent questions include the amount of data required for an optimal presentation. This may require further investigation into how users perceive auditory information to identify the amount of data which can be effectively encoded in an auditory signal to actually convey useful information to the user.

At the same time, multimodal displays should avoid auditory and visual overload when presenting data. Visual and auditory cues, such as visual modifications or spatial audio, can be used to direct the user's attention to a particular area; however, this requires the algorithm to understand which data is most relevant to the user. Effective use of audio-visual cues may also require identifying need-specific sonification algorithms or methods, as the same auditory display may not be equally effective in all tasks, or even subtasks. A particular sonification algorithm may be well suited for improving understanding of a dataset, but might be less useful in time-sensitive situations where real-time processing and decision making is required. These audio-visual cues can work together to avoid perceptual overload. For instance, certain sound patterns can have a calming effect on users' cognitive states. In this way, the delivery of auditory information may help focus the user's attention to reduce perceptual load and allow for better processing.

Audio-visual presentations should also maintain smooth transitions between deliveries of data. For instance, when the user navigates to a new geographical area on the virtual globe, they may experience a sudden change in the auditory display of data, which may sound incoherent or unintelligible. The challenge lies in identifying sonification algorithms which can effectively highlight the transition of data without alarming or disorienting the user. Providing coherent and informative transitions between auditory representations may be done through the use of cues such as spatial audio, which helps users understand spatial relationships between data.

A final observation is the need for identification of optimal sonification algorithms for enhancing the user's cognition of data in a variety of scopes or tasks. Avenues such as rule-based sonification may permit users to control the aural representation of data in greater detail. The application of interactive methods may allow users to actively identify and sonify data rather than simply perceive fixed aural representations of data. Interactivity should remain a driving force in the development of multimodal displays of data.

Our contribution presents an initial but essential step towards a further exploration of cross-modal spatial data infrastructures capable of supporting the user's exploration of data. Research opportunities can leverage the potential of these infrastructures and integrate them with advances in user interaction technology to allow richer explorations of spatial data through more user-centric presentations. These displays can augment user cognition of geospatial information, promote accessibility and explainability of spatial data, and help decision makers address the broader challenges humanity currently faces.

Bibliography

[1] F. Prandi, F. Devigili, M. Soave, U. Di Staso, and R. De Amicis. 3D web visualization of huge CityGML models. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 601–605, La Grande Motte, 2015. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

[2] Jae Gil Lee and Minseo Kang. Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, 2(2):74–81, 2015.

[3] David J. Maguire and Paul A. Longley. The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, 29:3–14, 2005.

[4] Thomas Blaschke, Karl Donert, Frank Gosette, Stefan Kienberger, Martin Marani, Qureshi Salman, and Dirk Tiede. Virtual Globes: Serving Science and Society. *Information 3*, 2012.

[5] Alan M. MacEachren and Menno-Jan Kraak. Research Challenges in Geovisualization. *Cartography and Geographic Information Science*, 28(1):3–12, 2001.

[6] Zhihan Lv, Xiaoming Li, Baoyun Zhang, Weixi Wang, Shengzhong Feng, and Jinxing Hu. Big City 3D Visual Analysis. In *Eurographics 2015*, pages 11–13. Eurographics, 2015.

[7] Hossein Mojaddadi, Biswajeet Pradhan, Haleh Nampak, Noordin Ahmad, and Abdul Halim bin Ghazali. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomatics, Natural Hazards and Risk*, 8(2):1080–1102, 2017.

[8] T. Vopham, J.E. Hart, F. Laden, and Y.-Y. Chiang. Emerging trends in geospatial artificial intelligence (geoAI): Potential applications for environmental epidemiology. *Environmental Health: A Global Access Science Source*, 17(1):1–6, 2018.

[9] Juan Sandino, Geoff Pegg, Felipe Gonzalez, and Grant Smith. Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors (Switzerland)*, 18(4):1–17, 2018.

[10] Georgios Kouziokas and Konstantinos Perakis. Decision Support System Based on Artificial Intelligence, GIS and Remote Sensing for Sustainable Public and Judicial Management. *European Journal of Sustainable Development*, 6(3):397–404, 2017.

[11] Rachel Robertson. Research aims to make artificial intelligence explain itself, 2017.

[12] Kristen Vincent, Robert E. Roth, Sarah A. Moore, Qunying Huang, Nick Lally, Carl M. Sack, Eric Nost, and Heather Rosenfeld. Improving spatial decision making using interactive maps: An empirical study on interface complexity and decision complexity in the North American hazardous waste trade. *Environment and Planning B: Urban Analytics and City Science*, 0(0):1–18, 2018.

[13] A. Marzouki, F. Lafrance, S. Daniel, and S. Mellouli. The relevance of geovisualization in Citizen Participation processes. In *Proceedings of the 18th Annual International Conference on Digital Government Research - dg.o '17*, pages 397–406, 2017.

[14] F. Prandi, G. Panizzoni, D. Magliocchetti, F. Devigili, and R. De Amicis. WebGL virtual globe for efficient forest production planning in mountainous area. In *Web3D '15: Proceedings of the 20th International Conference on 3D Web Technology*, pages 143–151, 2015.

[15] Tooba Nasir and Jonathan C. Roberts. Sonification of Spatial Data. In *Proceedings of the 13th Meeting of the International Conference on Auditory Display*, volume 1, pages 112–119. International Conference on Auditory Display, 2007.

[16] Sara Adhitya and Mika Kuuskankare. The Sonified Urban Masterplan (SUM) : Sonification For Urban Planning And Design. In *Proceedings of the 17th Meeting of the International Conference on Auditory Display*. International Conference on Auditory Display, 2011.

[17] Emma Chow, Amin Hammad, and Pierre Gauthier. Multi-touch Screens for Navigating 3D Virtual Environments in Participatory Urban Planning. In *Extended Abstracts on Human Factors in Computing Systems*, pages 2395–2400. SIGCHI, 2011.

[18] Jocelyn Parong and Richard E Mayer. Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology*, 2018.

[19] Gerhard Gröger, Thomas H Kolbe, Claus Nagel, and Karl-Heinz Häfele. OGC City Geography Markup Language (CityGML) Encoding Standard. Technical report, Open Geospatial Consortium, 2012.

[20] Arne Schilling, Jannes Bolling, and Claus Nagel. Using glTF for streaming CityGML 3D City Models. In *Proceedings of the 21st International Conference on Web3D Technology*, pages 109–116. Web3D, 2016.

[21] Patrick Cozzi. Introducing 3D Tiles. The Cesium Blog, 2015.

[22] Saurabh Bhatia, Patrick Cozzi, Alexey Knyazev, and Tony Parisi. glTF Runtime 3D Asset Delivery, 2017.

[23] Katalin Tóth, Clemens Portele, Andreas Illert, Michael Lutz, and Vanda Nunes de Lima. A Conceptual Model for Developing Interoperability Specifications in Spatial Data Infrastructures. Technical report, Joint Research Centre, European Commission, 2012.

[24] Yingjie Hu and Wenwen Li. Spatial Data Infrastructures. *The Geographic Information Science & Technology Body of Knowledge*, 2017.

[25] Mark Gahegan. Geovisualization. 2001.

[26] P.K Singh. Spatial Data Infrastructure in India : Status , Governance Challenges , and Strategies for Effective Functioning. *International Journal of Spatial Data Infrastructures Research*, 4:359–388, 2009.

[27] Ian Masser. All shapes and sizes: The first generation of national spatial data infrastructures. *International Journal of Geographical Information Science*, 13(1):67–84, 1999.

[28] Chaowei Yang, Robert Raskin, Michael Goodchild, and Mark Gahegan. Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems*, 34(4):264–277, 2010.

[29] Hossein Mohammadi, Abbas Rajabifard, and Ian Williamson. Spatial Data Integrability and Interoperability in the Context of SDI. In Lars Bernard, Hardy Pundt, and Anders Friis-Christensen, editors, *Lecture Notes in Geoinformation and Cartography*, chapter 4, pages 401–413. Springer-Verlag Berlin Heidelberg, 2008.

[30] C. Zhang, W. Li, M. J. Day, and Z. R. Peng. GML-based interoperable geographical databases. *Cartography*, 32(2):1–16, 2003.

[31] Martin Breunig and Sisi Zlatanova. 3D geo-database research: Retrospective and future directions. *Computers & Geosciences*, 37(7):791–803, 2011.

[32] PostGIS 2.4.3dev Manual. PostGIS, 2017.

[33] John E. Bailey and Aijun Chen. The role of Virtual Globes in geoscience. *Computers and Geosciences*, 37(1):1–2, 2011.

[34] G. Conti, D. Magliocchetti, F. Devigili, and R. De Amicis. The role of SDIs as enablers of next generation intelligent transport systems. In Otakar Cerba and Kristyna Cerbova, editors, *SDI, Communities and Social Media*, number September, pages 71–90. Czech Centre for Science and Society, 2012.

[35] Zhong Ren Peng and Chuanrong Zhang. The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS). *Journal of Geographical Systems*, 6(2):95–116, 2004.

[36] Thomas H Kolbe, Gerhard Gröger, and Lutz Plümer. CityGML: Interoperable Access to 3D City Models. In Peter van Oosterom, Siyka Zlatanova, and Elfriede M Fendel, editors, *Geo-information for Disaster Management*, pages 883–899. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[37] R. Esmaeili and F Karimipour. Cartographic representation of spatial data quality in VGI for users with different semantics. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(3W3):81–84, 2015.

[38] B. Simes, G. Conti, S. Piffer, and R. De Amicis. Enterprise-level architecture for interactive web-based 3D visualization of geo-referenced repositories. *Proceedings of Web3D 2009: The 14th International Conference on Web3D Technology*, pages 147–154, 2009.

[39] Thomas H Kolbe, Zhihang Yao, Claus Nagel, Richard Redweik, Philipp Wilkomm, Gyrgy Hudra, Arda Müftüoglu, and Felix Kunde. 3D City Database for CityGML. Technical University of Munich, 2016.

[40] Keith V. Nesbitt. A classification of multi-sensory metaphors for understanding abstract data in a virtual environment. In *2000 IEEE Conference on Information Visualization. An International Conference on Computer Visualization and Graphics*, pages 493–498. IEEE, 2000.

[41] Trevor Hogan and Eva Hornecker. How Does Representation Modality Affect User Experience of Data Artifacts? In *Proceedings of Haptic and Audio Interaction Design (HAID12)*, pages 141–145. Springer, 2012.

[42] Stephen Barrass, Jonathan Berger, Terri L. Bonebright, Till Bovermann, Eoin Brazil, Stephen Brewster, Densil Cabrera, Simon Carlile, Perry Cook, Alberto

de Campo, Florian Dombois, Gerhard Eckel, Alistair D. N. Edwards, Alfred Effenberg, Mikael Fernström, Sam Ferguson, John Flowers, Karmen Franinovic, Florian Grond, Anne Guillaume, Thomas Hermann, Oliver Höner, Andy Hunt, Gregory Kramer, Guillaume Lemaitre, William L. Martens, David McGookin, Michael Nees, John G. Neuhoff, Sandra Pauletto, Michal Rinott, Niklas Röber, Davide Rocchesso, Julian Rohrhuber, Stefania Serafin, Paul Vickers, and Bruce N. Walker. *The Sonification Handbook*. Open Access, 4 edition, 2011.