NASA WEBWORLDWIND:
MULTIDIMENSIONAL VIRTUAL GLOBE FOR GEO BIG DATA VISUALIZATION

M. A. Brovelli [a], P. Hogan [b], G. Prestifilippo [a]*, G. Zamboni [a]

[a] Politecnico di Milano, DICA, Laboratorio di Geomatica, Como Campus, via Valleggio 11, 22100 Como, Italy -
maria.brovelli@polimi.it, gabriele.prestifilippo@mail.polimi.it, giorgio.zamboni@polimi.it
[b] NASA Ames Research Center, M/S 244-14, Moffett Field, CA USA - patrick.hogan@nasa.gov

Commission II/ThS 12 - Location-based Social Media Data

KEY WORDS: Virtual Globe, Data Visualization, Big geo-data, Web GIS, Multi-dimensional data, Social Media

ABSTRACT:

In this paper, we presented a web application created using the NASA WebWorldWind framework. The application is capable of visualizing n-dimensional data using a Voxel model. In this case study, we handled social media data and Call Detailed Records (CDR) of telecommunication networks. These were retrieved from the "BigData Challenge 2015" of Telecom Italia. We focused on the visualization process for a suitable way to show this geo-data in a 3D environment, incorporating more than three dimensions. This engenders an interactive way to browse the data in their real context and understand them quickly. Users will be able to handle several varieties of data, import their dataset using a particular data structure, and then mash them up in the WebWorldWind virtual globe. A broad range of public use this tool for diverse purposes is possible, without much experience in the field, thanks to the intuitive user-interface of this web app.

1. INTRODUCTION

1.1 State of the Art

"Life happens in three dimensions, so why doesn't science?" (Butler, 2006). In recent decades, Virtual Globes have changed the way we interact with visualization of data, improving the user experience for understating different phenomena. Their proliferation followed the development of VRML - Virtual Reality Modeling Language - in 1994. This provided the opportunity to share a Virtual Globe on the web and permitted the web community to see and interact with a three-dimensional representation of Earth.

In those early years, developing with VRML was something reserved for a select few since it needed sophisticated programming skills to create simple applications. With the start of the new century and the increasing evolution of the Web, new technologies and improved tools made the 3D visualization more available to a much larger community. Earth Viewer, created by Keyhole Inc., was presented in 2001 and afterward acquired by Google in 2004. Earth Viewer was renamed to Google Earth in 2005 and became a point of reference for many scientists and the public at large.

Meanwhile, another important point of reference for the GIS community was the release in 2003 of NASA WorldWind as open source. With NASA WorldWind, the approach to virtual globes changed, greatly expanding the user community who could utilize Virtual Globes to interact with geo-data. The main difference WorldWind provided versus Google Earth was the opportunity to customize a virtual globe via fully customizable applications.

Even though this was just the beginning of the Virtual Globe era, new possibilities started to advance into broader areas of study. What created a favorable development and promotion of Virtual Globes was the increase of geospatial data available to use with them. With the advances in the Remote-Sensing field, and new data collected from satellites, aircraft, and various sensors, the need for a visualization tool became fundamental, and the use of Virtual Globes began to escalate in numerous sectors: Education, Research, Disaster Response, Data Analysis and verification, Geo-Collaboration and much more.

Michael F. Goodchild Professor of Geography at the University of California said: "It's like the effect of the personal computer in the 1970s, where previously there was quite an élite population of computer users" (M. Goodchild, 2006). The élite cadre of geographic information system (GIS) users is still rapidly increasing thanks to the advantage of Virtual Globes now running on the Web, without the need for any external components or installing of additional software.

1.2 Presentation structure

The paper presents some relevant tools developed with NASA Web WorldWind and is structured as described here. In the first section, some background on Virtual Globes is provided. We focused on Free and Open Source software packages. We are interested in developing advanced functionalities, and so we needed the complete freedom of working with an open platform. And we wanted an already existing platform since we had decided not to develop a Virtual Globe from scratch. We continue in this section with a short description of the approaches used in 2D and 3D for visualizing geospatial metadata. We then explained the voxel model, used in our case, with other examples of similar models, used in other studies.

Our case is then explained, indicating the sample dataset we considered, the possible approaches to similar kinds of data and the specific features implemented in our application. To conclude, we suggest possible future developments, to extend the application and adapt it to work with more and different datasets.

1.3 Virtual Globes

Two notable examples of modern Virtual Globes are NASA WebWorldWind and AGI Cesium. These run in any browser thanks to the support of JavaScript and HTML5 technologies. It means these online tools are cross-platform and can run on any

---

* Corresponding author

device (PC, Mac, Smartphone, Tablet, etc.). What is especially interesting about these new Virtual Globes is the possibility to customize them without having advanced programming knowledge skills. This means that any user can create a Virtual Globe application and easily share it with everyone via the web. Thanks to the web-capable Virtual Globe, dealing with three-dimensional data has become much more accessible.

WebWorldWind and Cesium, along with their APIs, Application Programming Interface, offer feasible ways to display a wide range of data.

WebWorldWind is designed to be easily extensible, thanks to an interface on top of JavaScript and WebGL. Due to this *geospatial* approach used in the WebWorldWind framework, any user with some experience in JavaScript can create components for it.

Even though WebWorldWind is at the dawn of its development life cycle, different organizations around the world are using it in many applications, such as Smart Cities, Terrain and City Visualization, Vehicle tracking, Geospatial data analysis and more.

2. VOXEL MODEL WITHIN VIRTUAL GLOBES

2.1 Models

Experiencing 3D data in its native context, inside a Virtual Globe, is fundamental in order to select the right model to show the multidimensional metadata. An appropriate model for this 3D environment is the voxel model; it consists of an array of voxels -akin to a cube- to show data with a volume representation.

Voxels have been used in many study areas; the Journal of Biology article "Computer Visualization of Three-Dimensional Image Data Using IMOD," voxels are used to view tomograms (Kremer, 1996). In "Beyond volume rendering: visualization, haptic exploration, and physical modeling of voxel-based objects," a voxel representation is proposed for modeling objects and structures derived from volumetric data (Gibson, 1995).

Several approaches based on voxels have been adopted to work with environmental data, in "Virtual globes for 4D environmental analysis" (M. A. Brovelli & G. Zamboni, 2012). In this work, using NASA WorldWind Java, there are several examples on how to handle environmental variables in multiple dimensions.

One example is a 4D visualization of Lake Como water temperature at different depths, represented with different colors. For the utilization of energy among the buildings in Olbia City, we represented the buildings by the color, and extruded the heights.

Another example developed in this paper is the implementation of the Environment Space and Time Web Analyzer (EST-WA) to manage environmental variables. In the sample, a 3D array of voxels is created from a NetCDF file and shown on the globe. Moreover, to view the inner voxels of the 3D object, the volume can be sectioned with planes orthogonal each other.

The aim of this paper is to show a reproduction of the work by M. A. Brovelli and G. Zamboni with the EST-WA in a web browser compatible Virtual Globe - WebWorldWind - and improve the previous product with additional features.
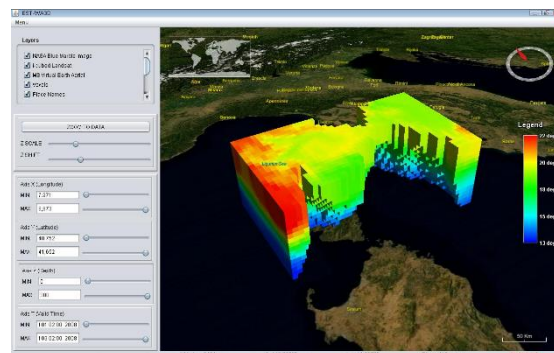

Figure 1. EST-WA 3D

Thus, we implemented a voxel model similar to EST-WA to show n-dimensional variables in a 3D environment. In our model, we represented the fourth dimension (3D space and the time) by the voxel's color. More possibilities are available by linking variable values to the opacity, outline stroke, shade, pattern and more. There are limits to a user's ability to quickly understand a model with many variables. As we can see in Figure 2 where we have a simple voxel with a variable linked to the color of the object, in contrast with a complex one, in Figure 3 where the outline color and a pattern in the shade may represent other two variables.
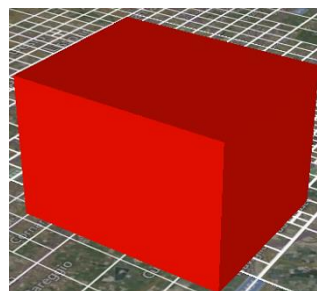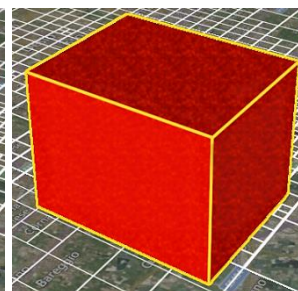

Figure 2. Simple Voxel   Figure 3. Complex Voxel

2.2 Application Architecture

The application is designed without the use of any external framework but WebWorldWind (https://webworldwind.org/).

As an external library to interact with the user interface of the application, we used jQuery (https://jquery.com/). To represent the elements in the web interface we used jQuery UI (https://jqueryui.com/) and Bootstrap (http://getbootstrap.com/). The architecture of the application is considered to be Object-Oriented. This kind of architecture is made of components that do not have strict dependencies on the other components; meaning that any part of the application can be re-used.

For instance, we can take into consideration the implementation of voxels: a voxel is an object generated by WebWorldWind functionalities. Anyone needing to implement a voxel model can easily copy the related file and import it into their application.
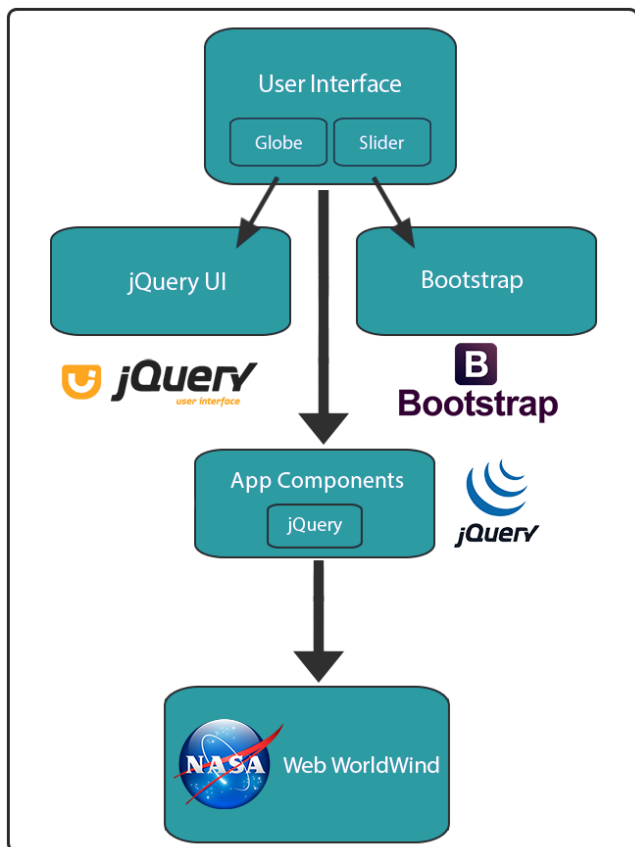
Figure 4. Application Architecture

3. CASE STUDY: TELECOMMUNICATION BIG DATA

In our study case, we wanted to address the visualization to telecommunication and social media data. "Global telecommunication services create an enormous volume of real-time data" (Koutsofios, 1999). Thus, visualizing these data is becoming a significant challenge, especially when we have several variables to represent. The use of 3D visualization can ease the user's understanding and interpretation of the data.

We took into consideration an important dataset: "Big Data Challenge 2015" from Telecom Italia. The data provided show information about telecommunication events, i.e., SMSs, calls and internet usage. All entries have a timestamp expressed in milliseconds, and each one represents information for 15 minutes from that timestamp.

The data refer to the year 2014, and we have all the 12 months of data for that year. In particular, for the SMSs, we have information about the number of received SMSs and sent SMSs, for the call we have the number of incoming calls and outgoing calls.

Regarding the internet data, we have the number of started connections, closed connections and amount of data transferred. The social media dataset contains geo-localized tweets originated from Milano. Each entry has an anonymous user identifier; thus, different entries could have a repeated user-id in the case of repeated tweets.

Telecom Italia provided a geo-referenced irregular grid, for some cities in Italy, where an id represents each grid cell. An irregular grid means that each grid element has a different size. The concept of rows and columns for regular grids does not apply anymore, and more peculiar is the fact to have a grid not circumscribed by a rectangular shape, but open to any geometric contour. In our case, the grid is made of cells varying from a

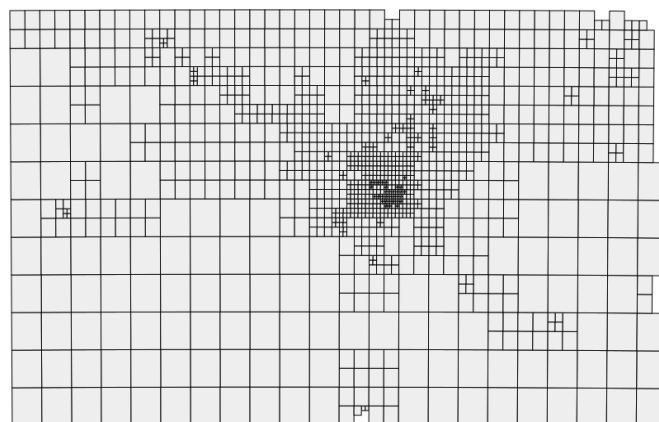dimension of 255x325 meters to 4080x5200 meters as shown in Figure 5.

Figure 5. Milano Grid, BigData Challenge - Telecom Italia

In the data, each entry has a grid-id number to link the value of the data to the corresponding grid cell, a timestamp, and value for each of the mentioned data. A singular fact is that there are several "holes"; this means some data are missing. In particular, in each timestamp, some values for the grid elements are not present.

The case study presented considers the city of Milano, and we show data about call-in and call-out. However, since all of the data has the same structure and refer to the same grid, we could work with all of them in the same way.

Using WebWorldWind, we thus defined a voxel model, which could place a volume representation in any position of the grid, taking as edges the bounds of each grid element. As we mentioned before, the data have no information on the altitude, and we could exploit this to represent time using altitude, the z-axis, permitting the users to have in view an accessible three-dimensional representation of the data within a time range. In this way, four dimensions are represented for a specific meta-data: value, latitude, longitude, and time. The same approach was adopted by M. A. Brovelli & G. Zamboni (2012).
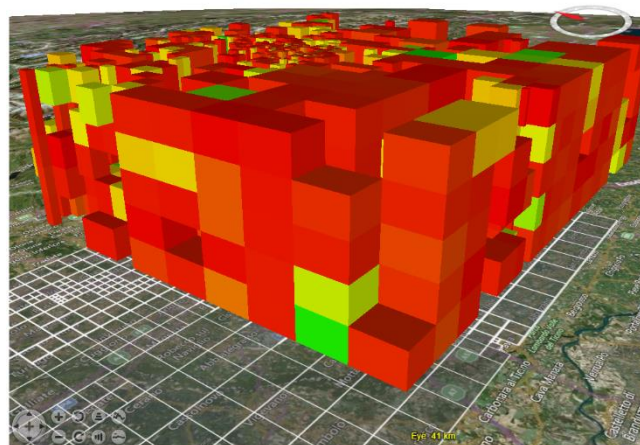

Figure 6. Voxels representing outgoing calls over Milano grid

In some cases, other researchers have used some simplification of the voxel model, to reduce the complexity. With a fully populated array of voxels, it is possible to show the data only on the external surface and not render all the voxels in the interior of the surface. What comes new in this case is the irregular grid and non-continuous entries for each timestamp.

WebWorldWind offers good performance in representing a significant amount the voxels, without the need for any

simplification. In this case, we created a way to navigate and observe the data, filtering out voxels within a range of three axes. Thanks to sliders in the application interface, we can limit the extension of latitude, longitude, and time that we want to show on the globe. The user has the ability to reveal the interior of the model using three cutting planes. Thus, we can observe fewer voxels easily reducing the complexity of the model.

We adopted a heuristic approach to show a statistical representation of data, using weighted average, variance, median, maximum and minimum. We allow users to subdivide the data into subsets, choosing a number of groups to present the data, with grouping possible for the altitude, longitude and time. This technique consists of a pyramid representation of voxels that permits showing of some statistics, alongside a larger amount of entries in time (see Figure 7).
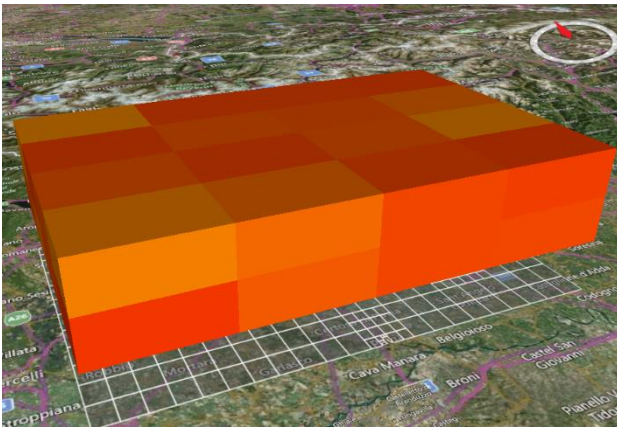

Figure 7. Group representation of average outgoing calls

Moreover, performance improvements came in handy as well. Having a lower number of voxels to represent - putting them in subgroups - allows better performances in the virtual globe. Thanks to this, any user can interact with the globe and the data. It is also possible to select a single voxel or a subgroup and obtain information regarding the available metadata about that voxel through the time in a simple 2D graph.

For as much data as we represented with altitude for time, not many layers of voxels can be incorporated because inserting more than a few, may be difficult to observe and also cause some slow-down in the globe. To solve this issue, we considered another way to navigate through time. We allow users to keep visible a specific number of time layers for a fixed time range, but the application still permits them to move this range to animate and view the desired time range.

We demonstrated the introduction of another way to show the time, but this technique can be used just as well to handle another variable in the scene if needed. Another feature we implemented consists in the possibility of comparing two different variables from a dataset or even from two different datasets within a single view.

We can imagine this type of comparison by creating a first representation of a certain variable, then creating another in a second Virtual Globe and then putting the two views side by side. Using another approach in order to see everything clearly, and to not complicate the user experience, we keep a single virtual globe. This approach allows selecting two variables and representing them in a single voxel, subdividing it into two parts. In this new voxel, we fill half the volume with the corresponding color value of the color scale for the first variable, while the remaining half with the color for the second variable. Having two values information on a single voxel is thus possible to retrieve the information about the two datasets in time alongside some statistics, as their correlation.

To improve the perception of the data inside a grid array we created also a slider to change the opacity of the voxels to appreciate the voxels hidden at first from the outside and reveal the surface of the underlying terrain. Moreover, filtering functionalities are provided in the application interface to also filter the data on their value, restricting it, so the range on the color scale shows a specific subset of values.

4. CONCLUSIONS AND FUTURE DEVELOPMENTS

The application developed satisfies the expectations for the visualization features and performance. Although we experienced some performance issues when visualizing a large amount of data, with the simplification techniques we applied, we can still process large quantities of data without performance issues. In our case study, the application shows data from the "Telecom Italia Big Data challenge" dataset, but we can visualize any data source in this format. However, we want users to import their dataset in any format. The application is already predisposed to any source of data; the only requirement would be to edit the data format importing system.

REFERENCES

Brovelli, M.A., Zamboni, G., 2012. Virtual globes for 4D environmental analysis. *Applied Geomatics* 4, pp. 163–172.

Brovelli, M. A., Zamboni, G., 2014. Environmental Space and Time Web Analyzer. *Proceedings of the 2014 Conference on Big Data from Space (BiDS'14), Research, Technology and Innovation (RT&I), Frascati, November 1214*, pp. 224–226.

Butler, D., 2006. *Virtual globes:* The web-wide world. *Nature* 439, pp. 776–778.

Gibson, S.F.F., 1995. Beyond volume rendering: visualization, haptic exploration, and physical modeling of voxel-based objects. *Eurographics Visualization in Scientific Computing '95* pp. 10–24.

Göbel, S., Jasnoch, U., 2001. Visualization techniques in metadata information systems for geospatial data. *Advances in Environmental Research* 5, pp. 415–424.

Jacoby, W.G., 1998. Statistical Graphics for Visualizing Multivariate Data.

Koutsofios, E., North, S., Keim, D., 1999. Visualizing large telecommunication data sets. *IEEE Computer Graphics and Applications* 19, pp. 16–19.

Kremer, J.R., Mastronarde, D.N., Mcintosh, J., 1996. Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology* 116, pp. 71–76.

Peuquet, D.J., Kraak, M.-J., 2002. Geobrowsing: Creative Thinking and Knowledge Discovery Using Geographic Visualization. *Information Visualization* 1, pp. 80–91.

Wootton, C., 2007. *Developing quality metadata: building innovative tools and workflow solutions*. Focal Press, Amsterdam