Optimized design of customized KML files*

Marianna Zichar

University of Debrecen
zichar.marianna@inf.unideb.hu

Abstract

The need of data visualization appears regularly in several disciplines. If the data has spatial component its visualization can be based upon it, so the need of using a simple but powerful tool for supporting the geovisualization is evident. There are several methods to create and represent such files, from which the usage of KML files are said to be the most popular [1]. Although the user interface of Google Earth (the native browser and producer of KML files) provides a number of tools for defining elements and saving them in kml or kmz format [2], these files cannot be regarded optimal. Several inconvenient situations could be avoided if the users followed simple guidelines and edited the generated files. This paper focuses on defining the main steps of reorganizing a KML file and provides designing instructions to get KML files with optimized structure.

Keywords: KML, customized content, geovisualization, Google Earth

MSC: 68U35, 68P20

1. Introduction

Data can appear in several forms depending on the purpose of their usage. Data visualization is a special way to represent data in graphical form emphasizing some characteristics playing important roles for us. Visual perception of humans may have rather powerful capabilities to synthesize information, seek patterns, and reveal hidden relationships between the objects if the data are represented visually. The same amount of data can be visualized in different ways, of course. A branch of the visualization techniques is the geovisualization which integrates approaches from visualization in scientific computing, cartography, image analysis, information visualization, exploratory data analysis, and geographic information systems to provide theory, methods and tools for visual exploration, analysis, synthesis and presentation of geospatial data. This definition comes from the Commission on

*The publication was supported in part by the TÁMOP-4.2.2C-11/1/KONV-2012-0001 project supported by the European Union, co-financed by the European Social Fund.

Visualization and Virtual Environments of the International Cartographic Association (ICA) [3].

About 80% of all digital data generated nowadays possess information which is interpreted in a geographical reference system. Due to this fact it is not surprising that the need of representing customized data based on their geospatial information has been appearing in many fields (e.g. archaeology, environmental studies, urban planning, medicine, linguistics, biology, agriculture) and researches from these disciplines are interested in geovisualizing their data in a convenient and costless way.

1.1. Keyhole Markup Language

Geovisualization requires maps and data in an appropriate format. If data analysis is out of our interest, then we do not need a complete geographical information system; we can realize our objectives by using the services of an online map provider. There are several companies offering such free services, but services of Google Inc. are probably the most popular. The Google Maps (GM) and the Google Earth (GE) are well-known applications even for the public, and their professional usage is getting to be frequent too. Both services require an internet connection, are embeddable into webpages, but Google Earth has a desktop version too, which can be used to create KML files containing data to visualize. The user interface of desktop Google Earth has a number of tools enabling the definition of KML features and also makes possible to save our data in form of KML or KMZ.

Keyhole Markup Language is an XML notation for describing information in a geographic context. The KML Version 2.2 became an international standard since it has been adopted as an OGC (Open Geospatial Consortium) implementation standard for presenting geographic information visually. Basically, the file extension `kml` refers to a simple, human-readable file which is composed of text and punctuation. It can be created and edited with a basic text editor, although the most typical way of creation are exporting data to KML or saving the features with the help of the graphical user interface after creating them with its tools (Figure 1.). A detailed overview of the creation methods is available in [4]. Publishing of KML files is also easy. Either you share the file itself (delegating the visualization to the users) or you create a webpage to host an embedded Google Maps or Google Earth. Although the last options require some HTML and JavaScript knowledge, it is not hard to set up a simple webpage thanks to the several tutorials available on the internet.

The KML community is wide and varied; more and more scientists, teachers, students, average people prefer visualizing their geospatial data in the form of KML/KMZ files. Usage of the graphical user interface of Google Earth is rather attractive and makes the user feel that he does not need to be a technical wizard to master the basics of KML. This is definitely true if the user do not need to go over a limit. Unfortunately, there are some KML elements (for example screen overly, multigeometry, region) that cannot be created via the user interface. The file size saved in Google Earth is not optimal, and the window Properties of placemarks

Figure 1: Tools for creating the basic types of placemarks

is not suitable to construct arbitrary customized descriptions. In the case of large amount of data, the performance of our system can decrease. We often need to edit the KML files manually, which requires extra efforts from the users mainly, if he is not experienced KML user.

2. Strategies for optimal design of KML files

As I summarized in the last section we cannot avoid of editing KML files and having some advanced knowledge about KML if we would like to create a powerful presentation of our data. My experiences show that average people face lots of challenges during creating KML files after a while, because it does require sophisticated knowledge and know-how to overcome the issue. This section gives an overview about the issues and their solutions arising during using the user interface of GE to create KML files.

2.1. Customization

First, I provide a brief overview what can be meant under customization of a KML file. The first things, which come to the mind of the user when mentioning the word customization are the design of the bubble content and the graphical appearance of the features. Both activities are supported by the user interface but follow-up editing may be needed. The graphical appearance of the features is determined by styling independently on the geometry type. The user interface of GE generates usually redundant styling (increasing the file size), because shared styling is applied only to features belonging to the same folder. The default appearance of the bubble contains the value of the following KML elements: name and description (and two links for driving directions: To here – From here). The element description can contain plain text, as well as HTML and JavaScript.

Usage of advanced features, enhance the performance, but not necessary available from the user interface. To use some of more advanced features (such as region, screen overlay, timespan/ timestamp) the file has to be edited by the author, requiring extra knowledge.

Figure 2: The default appearance of a placemark
and a customized one

<https://sites.google.com/site/fusiontablestalks/stories>

2.2. Validation

There are applications which can be used to validate the content of a KML file. They can be a great help to check the file structure, to find errors in a file, and some of them are also able to define recommendations, suggestions to improve the quality. A document is said to be well-formed in XML if it adheres to the syntax rules specified by the XML 1.0 specification that is to say it satisfies both physical and logical structures.

jEdit is an open-source, Java-based text editor that works on several platforms (Windows, Mac OS, and Linux). There is an additional plug-ins for jEdit, which makes it possible to check whether a KML file is a well-formed document and is valid against the official KML schema. Basically, this application can help to reveal only syntax errors, which sometimes is a great challenge.

The Galdos KML Validator (www.kmlvalidator.com) is a more sophisticated, free, online service that can be used to validate KML files more strictly. Not only KML files but also KMZ archives can be submitted for validation through its simple web interface. Unfortunately, the online interface is configured to validate files up to 2 megabytes in size. Files larger than 2 megabytes, including KMZ archives containing KML files that are bigger than this size limit, will be rejected. That is why we should take care of the file size and pay attention to my recommendations described in the next section. The purpose of the KML Validator is to assist KML data authors in producing the highest quality KML by highlighting where their KML files are invalid or where they could be improved (Figure 3).

2.3. Optimization

The objectives of optimization are to avoid redundancy and slow rendering and finally to improve the performance of our KML file(s). To reach all of these objectives it is necessary to get to know the KML deeper and to edit the file(s). The online references, tutorials and also the traditional books describe the function-

Figure 3: The report after validating an invalid KML file

ality of the elements, but do not define strategies to solve practical issues [5]. I contacted several researchers from different disciplines who missed these guidelines, so I collected my recommendations.

The file size can be critical in many aspects. Large files, for example, cannot be validated, are rendered slowly, can be overcrowded. The code of placemarks saved by GE contains some elements which could be ignored in most cases. My suggestion is to remove the element LookAt and also the non-standard element with the prefix rx: . Figure 4 displays that the LookAt element itself consists of nine lines and that the gx:drawOrder element is included in the code of a placemark automatically, although in the case of points this element can be omitted. Another possibility to reduce the file size to have GE reused shared styles. Every placemark design is determined by a style map. Even the most simple style map requires 28 lines of code. It is worth to consider whether we need stylemaps indeed. The default stylemap is working without explicit declaration too. If placemarks are grouped into folders during their preparation, GE tries to use the same style map definition in the case of the same placemark style. It is important to remark, that moving a placemark into a folder subsequently does not result simplification in styling. Inline formatting is not used by GE at all, so this can be another option to spare lines when only little modification is needed in styling. To avoid too much editing a Python script can be written to automatize the tasks.

```
<Placemark>
  <name>Sample placemark</name>
  <description>GE assigns the LookAt element to every placemark.</description>
  <LookAt>
    <longitude>21.65274635423911</longitude>
    <latitude>47.52433743738901</latitude>
    <altitude>0</altitude>
    <heading>1.226974428330999</heading>
    <tilt>6.865910609198387</tilt>
    <range>7326.594403656623</range>
    <gx:altitudeMode>relativeToSeaFloor</gx:altitudeMode>
  </LookAt>
  <styleUrl>#m_ylw-pushpin</styleUrl>
  <Point>
    <gx:drawOrder>1</gx:drawOrder>
    <coordinates>21.64840012021328,47.52672120830839,0</coordinates>
  </Point>
</Placemark>
```

Figure 4: Code of the simplest placemark

If the previous methods could not reduce the file size under the required level, then the features have to be distributed into separate files. At the same time a master KML is created referencing the files in a NetworkLink element. It is recommended to avoid absolute referencing of the files, and not to forget that GE does not add the referenced files to the archive when saving the master file in KMZ format.

3. Conclusions

To capture the information content of data can be easier if we visualize them in an appropriate way, because human understanding, exploring and analyzing of (large) data sets can be enhanced with visualization. This paper explored that although the tools of Google Earth user interface provide a convenient way to create KML files, it may be not efficient enough. I highlighted why the file size is important and what strategies have to be followed to reduce the file size. To have optimal KML files we cannot avoid editing, which requires deeper KML knowledge.

References

- [1] BALLAGH, L.M., RAUP, B. H., DUERR, R.E., KHALSA, S. J. S., HELM, C., FOWLER, D., GUPTE, A., Representing data sets in KML: Methods and challenges, *Computers and Geosciences*, Vol. 37 (2011), 57–64.
- [2] GOODCHILD, M. F., The use cases of digital earth, *International Journal of Digital Earth*, Vol. 1 (1) (2008), 31–42.
- [3] MACEACHREN, A. M., KRAAK, M. J., Research challenges in geovisualization, *Cartography and Geographic Information Science*, Vol. 28(1) (2001), 3–12.
- [4] ZICHAR, M., Geovisualization based upon KML, *Journal of Agricultural Informatics* Vol 3, No 1. (2002), 19-26.
- [5] WERNECKE, J., *The KML Handbook*, Addison-Wesley, 2009.