Big Data Visualisation with Fractal Geometry

James Rosindell
Imperial College London
Silwood Park Campus, Buckhurst Road, Ascot, Berks SL5 7PY, UK
j.rosindell@imperial.ac.uk

1. INTRODUCTION

Huge quantities of data are now routinely created and stored in science and industry. This revolution has led to a new challenge: how can we intuitively explore and visualise such a wealth of information? We will discuss the use of fractal-based algorithms combined with a deep-zooming user interface as a new method for solving the problem. To demonstrate the power of this approach, we focus on one specific example: visualising the tree of life, an evolutionary tree showing the diversity of life and the evolutionary links between all species.

2. BENEFITS OF A ZOOMING USER INTERFACE

It is striking, considering the flexibility of modern computing devices, how our data exploration remains tied to the old fashioned 'paper paradigm'. Whilst the data we attempt to visualise would be impossible to produce, store and transmit in paper form, we still view web *pages* and *scroll* up and down content that would appear similar if it were printed. We can do much more than this in the digital world and so one would expect that the optimal solution to any big data visualisation problem would escape the paper paradigm and make use of this freedom to create novel solutions.

We have been navigating our way around big data for thousands of years: the world around us is complex and filled with overwhelming quantities of information. We deal with this by getting an overview of our surroundings with the choice to focus on, and move closer to, any objects of interest in order to find out more about them. A deep-zooming interface to a hierarchically organised visualisation of data provides an intuitive way to mimic this experience: zooming in to drill down into the data and reveal relevant details.

3. FRACTAL VISUALISATION APPLICATIONS

A specific example that we will focus on as a proof of concept is visualisation of the tree of life. Such trees are a problem for visualisation because there

is no natural coordinate system to map each node and leaf of the tree to a position in space, classic methods become unwieldy and difficult to explore for trees with just one thousand leaves whilst the complete tree of life has millions of leaves. Furthermore, there is metadata associated with the leaves and nodes of the tree including graphs, maps and images. To solve this problem, we use fractal algorithms based on L-systems to lay out an appropriate tree structure on the page at different scales producing suitable objects for the zooming interface (Figure 1). These algorithms produce forms reminiscent of live trees providing a powerful visual appeal that is attractive to non-experts.

Figure 1: A composite of screen shots from OneZoom showing the location of humans nested in the tree of life

We plan to apply the same methods in the future to visualise other data. For example, directory structures, genealogies, industrial plant monitoring systems, decision trees and even source code.

4. REFERENCES

- Rosindell, J., and Harmon, L. (2012) OneZoom: A Fractal Explorer for the Tree of Life. *PLoS Biology*, 10(10), e1001406.
- Rosindell, J. *et al.* (2012) www.onezoom.org (retrieved 21 March 2014).