A 3D Treemap Approach for Analyzing the Classificatory Distribution in Patent Portfolios

Mark Giereth Harald Bosch Thomas Ertl

Visualization and Interactive Systems Institute (VIS), University of Stuttgart, Germany

ABSTRACT

Due to the complexity of the patent domain and the huge amount of data, advanced interactive visual techniques are needed to support the analysis of large patent collections and portfolios. In this paper we present a new approach for visualizing the classificatory distribution of patent collections among the International Patent Classification (IPC) – today's most important internationally agreed patent classification system with about 70.000 categories. Our approach is based on an interactive three-dimensional treemap overlaid with adjacency edge bundles.

Keywords: Patent Co-Classification Analysis, 3D-Treemaps, 3D Edge-Bundling

1 INTRODUCTION

Patent classification systems intellectually organize the huge number of patents into pre-defined technology classes. Patent classification is based on the outcomes of investments of the patent offices. Due to the diversity of today's technological fields, a patent may be classified with several classes. Technology classes could therefore be an appropriate unit of co-classification analysis to show linkages between technical fields.

The aim of our approach is to improve analysis of large patent collections and portfolios by visualizing their classificatory distributions amongst the International Patent Classification (IPC) [1], today's most important internationally agreed patent classification system with more then 70.000 categories. The original purpose of the IPC was to be an effective search tool for the retrieval of patent documents [2].

But the IPC also became an important source for investigating the state of the art in given fields of technology and the assessment of technological development in these fields. The IPC further serves as taxonomy when comparing the patent portfolios of different applicants. This latter aspect is the focus of this poster.

2 IPC STRUCTURE

The IPC is divided into a *core* and an *advanced* level. The core level is intended for general information purposes and for smaller, national patent collections, whereas the advanced level is intended for searching larger, international patent collections. The more detailed subdivisions of the advanced level are compatible with the core level.

The IPC is hierarchically organized into currently 8 sections, 129 classes, 633/639 subclasses (core/advanced), 7,066/7,315 main-groups (core/advanced), and 10,047/61,403 subgroups (core/advanced).

{giereth|bosch|ertl}@vis.uni-stuttgart.de

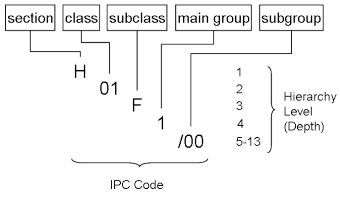

Figure 1.IPC code structure

We refer to elements of the IPC independently of their hierarchy level as *categories* to distinguish from classes as 2nd level elements. Each IPC category has a code, a label, and further attributes, such as version, 'see-also'-references to other categories, or associated keywords.

2.1 Co-Classification

A patent can be classified with more then one IPC category. There are two basic types of classification: a) classified as invention-information and b) classified as additional-information.

The classification begins at main group level. Top-level categories are never classified directly. This is the reason why we need a mechanism that allows top-level categories to 'inherit' from deeper categories by summing up their patents. As a further requirement it should be possible to compare the IP portfolio of two or more applicants. As a consequence, each category has as two associated patent sets for each applicant: directly classified patents and inherited ones.

Figure 2 shows a simple co-classification example, where a patent (EP0095852A1) is classified as G11B7/08, G11B7/085 and G11B7/09. This results in the three co-classification relations between these categories.

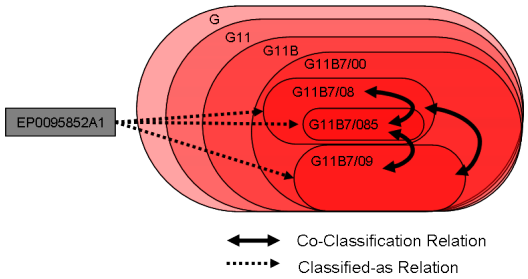

Figure 2.Example for patent co-classification

3 VISUALIZATION METHOD

To visualize the classificatory distribution of one or more patent portfolios, we generated an interactive 3D treemap, in which the third dimension represents the number of patents associated with a category. In contrast to other approaches [3,4] we synchronized a 2D and 3D treemap layout, where the 2D layout is used for generating the textures for the 3D items. This approach is briefly described in the following section.

3.1 3D Treemaps with 2D Textures

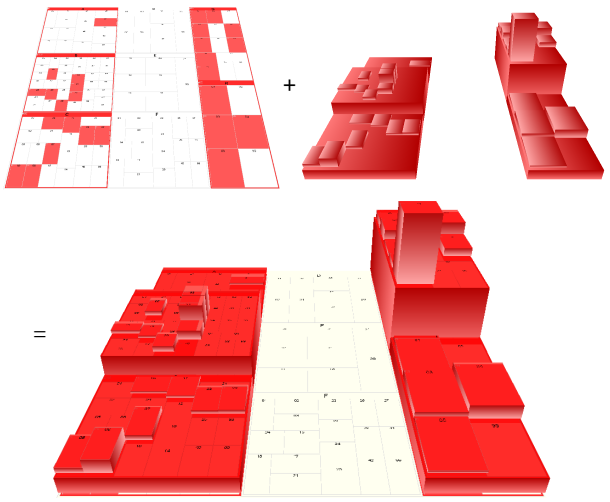

Figure 3.Composition of 3D treemap with 2D treemap textures

The 3D treemap visualization is based on the ordered treemap algorithm described in [5], which creates a sorted 2D layout for the IPC hierarchy. The number of patents associated with each category is then mapped to the 3rd dimension forming 'category cubes'. To allow a more efficient rendering, categories that do not have associated patents are represented by a 2D texture. Fig. 3 shows the composition for the 3D treemap for the IPC hierarchy up to the 2nd level. Each leaf category has a uniform area of 1.

3.2 Stacked 3D Treemaps

Multiple patent sets per category are stacked upon each other like in a stacked-bar-chart. The result is shown in figure 4. The texture on top of the stack is colored according to the larger stack (either blue or red in the example).

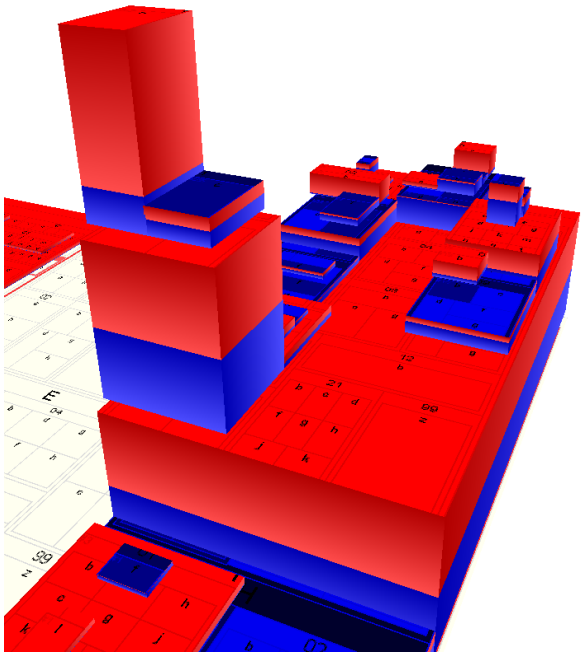

Figure 4.Stacked 3D treemap for classification analysis of two patent sets of different applicants (shown in red and blue)

3.3 3D Edge Bundling

The co-classification relations are visualized as 3D edge bundles, which use the hierarchic edge bundling algorithm presented in [6] in an extended version for 3D. The direct implementation of the bundling algorithm has the consequence that the resulting bundles go through the category cubes. Therefore they are rendered transparently. In figure 5 the co-classification relations of a single patent set are shown, where the right upper category has been selected. The selected edge bundles are shown in red, the others in blue.

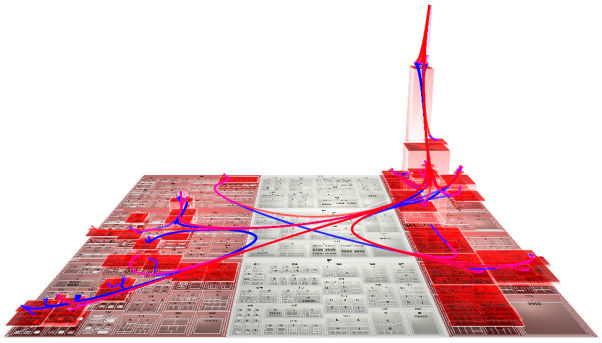

Figure 5.Co-classification relations shown as 3D edge bundles

4 DISCUSSION AND CONCLUSIONS

For implementing the prototype we extended the *prefuse* framework [7] with a new 3D treemap layout, a 3D edge bundling algorithm, and new OpenGL rendering components using the Java Binding for the OpenGL API (JOGL). We attached the visualization frontend to a patent retrieval backend, that both have been developed within the PatExpert project [8] funded by the European Commission.

The proposed visualization technique combines a 3D treemap layout with 2D treemap textures. The visualization aims in supporting the co-classification analysis of large patent sets – typically as the result of a previous patent query – by visualizing co-classification relations as 3D edge bundles. It also allows the comparison of patent portfolios of different applicants by utilizing 'stacked category cubes'. Stacked cubes could also be used for analyzing the temporal development of categories. Here, each stack would represent a time period (e.g., a year) indicated by a specific color. This latter aspect is subject to future work.

REFERENCES

[1] IPC homepage at the World Intellectual Property Organization (WIPO), http://www.wipo.int/classifications/ipc/en/

[2] A Bruun: Development of the IPC as a Search Tool, In *World Patent Information* (WPI), 21 (2):97-100, 1999

[3] T. Bladh, D. Carr and J. Scholl: Extending tree-maps to three dimensions: a comparative study, In proc. of 6th Asia-Pacific Conference on Computer-Human Interaction, pages 50-59, 2004

[4] H.-J. Schulz, M. Luboschik, H. Schumann: Exploration of the 3D Treemap Design Space, Poster at the IEEE Information Visualization Conference, 2007

[5] B. Shneiderman, M. Wattenberg: Ordered Treemap Layouts, In proc of *IEEE Symposium on Information Visualization*, 2001

[6] D. Holten: Hierarchic Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data, *IEEE Transactions on Visualization and Computer Graphics 12(5)*, 2005

[7] J. Heer, S. Card, and J. Landay: Prefuse: A Toolkit for Interactive Information Visualization, In proc. of ACM Human Factors in Computing Systems (CHI), 2005

[8] PatExpert project homepage: http://www.patexpert.org