Visualizing Metadata: LevelTable vs. GranularityTable in the SuperTable/Scatterplot Framework

Tobias Limbach, Peter Klein, Frank Müller, Harald Reiterer

University of Konstanz
Universitätsstr. 10, Box D73, 78457 Konstanz, Germany
{Tobias.Limbach, Peter.Klein, Frank.Müller, Harald.Reiterer}@uni-konstanz.de

Abstract

This paper will describe ongoing efforts to ease the information retrieval process on metadata, using a SuperTable/Scatterplot framework recently named *VisMeB* (**Vis**ual **Me**tadata **B**rowser). Based on the combination of a SuperTable (in two design variants) and a Scatterplot (also in two design variants), users can follow different search strategies to achieve results. Usability testing has showen, that an integration of our proposed visualizations supports most user's search style.

1 Introduction

To visually explore large data bases, the generation of metadata and their meaningful visualization is common practice. Systems like FilmFinder (Ahlberg & Shneiderman 1994) and xFIND (Andrews Gütl, Moser, Sabol, & Lackner 2001) not only enable users to search with keywords, they also let the users browse through result sets and enable a user based visual assessment of the information retrieval process. Users can decide, based on their estimation of the individual relevance of each iteratively retrieved result set, if their informational need has been satisfied or not. But, to put users in the position to decide, the system has to be usable, task oriented and support the users needs in each step of the information retrieval process.
After this introduction a brief overview of the system VisMeB is given, followed in the third chapter by an overview of usability test results of the two integrated design variants, while chapter four gives an outlook to current and future work.

2 Visualizing Metadata in a SuperTable/Scatterplot Framework

Based on a redesign of the Web search system INSYDER (Reiterer, Mußler & Mann 2001), we developed a visual metadata browser with different strategies to avoid the cognitively demanding traditional representation of metadata result sets, which often leads to a long list of document attributes. We assume that users don't want to scroll through endless result lists without any possibility to filter or sort out, the way they usually would have to, following the traditional concept of search engines.
Visualizations like a ResultTable, a BarGraph, and Tilebars, which were until then used separately, were put into one embedding visualization called SuperTable complemented by a Scatterplot. This multifocal approach using focus-plus-context techniques is found in other table based visualizations, for example the TableLens (Rao & Card 1994). Users are offered different brushing and linking techniques between the two components. This was inspired by the work of North and Shneiderman (North & Shneiderman 2000). The Scatterplot, as a graphical

representation of the result set, gives a quick overview and defines own views by zooming, selecting and filtering to reduce the amount of hits, while the SuperTable follows the concept of a distortion-based table with BarGraphs to visualize document relevance (Veerasamy & Navathe 1995), and Tilebars (Hearst 1995). A RelevanceCurve is used to assess text segments, as their individual relevance is visualized through stacked columns. Documents can be analyzed in high detail. Our aim was to avoid the use of cognitively demanding result lists as long as possible. With the interaction possibilities the user is given, it should be possible to focus on the cognitive and intellectual resources only on the original document.

Following the scenario based design approach proposed by (Rosson & Carroll 2001) we developed different information retrieval process scenarios. While developing those, a second design variant for the SuperTable emerged from discussions with M. Eibl (Klein, Müller, Reiterer & Eibl 2002) and was built into the scenarios. Though it became necessary to specify them as *LevelTable* and, the newer one, as *GranularityTable*.

The first design variation, called LevelTable, was strongly influenced by lessons learned from the INSYDER approach. Metadata can be explored in four different views representing the "level of detail" for documents. We call this method of looking closer and closer at the details of a document the "focus of interest". The more one wants to know about a document, the deeper one has to look and the higher the level has to be, ranging from a mere graphical overview in the first level to the document as stored in the database in the last level. By doing so it combines several visualizations in one enclosing table to ease the interpretation and show their conceptual connections.

The second design variant follows a granularity concept, and with that further smoothens the change of the visualizations between the exploratory steps. The information retrieval process is ergonomically integrated into the semantic structure of the visualizations, as described by (Eibl 2002). Granularity is a term used in photography to describe the accuracy of pictorial presentations in a picture. The higher the granularity, the more details can be seen on a picture.

This idea is transferred to integrate the visualizations of the SuperTable smoothly: The lowest granularity results in a simple histogram, which states the overall importance of the single documents. The highest granularity would lead to a representation of the actual text. In between these margins are as many intermediate steps as are required to give the impression of a smooth transition (as far as possible technically). A visual and conceptual clue is given by a slider (known from e.g. media players, household appliances etc.). Figure 1 shows two screenshots of the html-mockups.

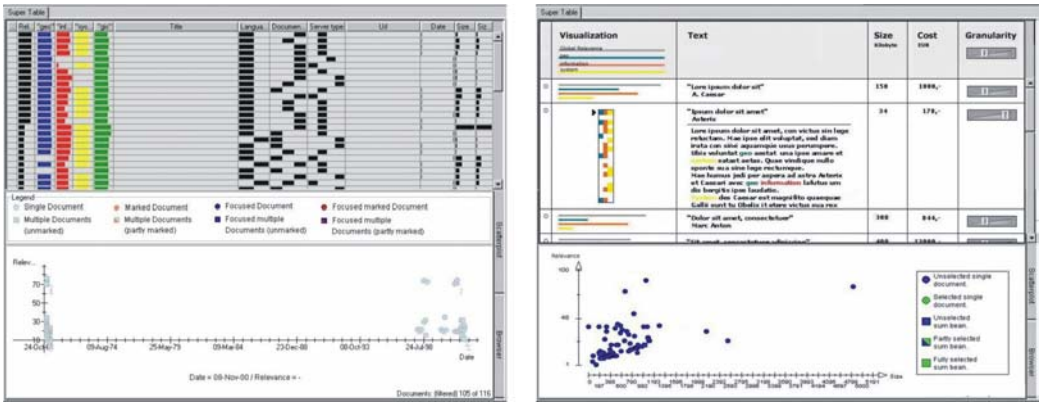

Figure 1: Html-mockups: LevelTable with overview, GranularityTable with different grades of detail

Both design variants are combined with a scatterplot. For early user tests, they were implemented as html-mockups with a limited range of interaction possibilities. Nevertheless, they fulfilled the requirements of our scenarios and were though tested.
A fundamental difference is established in the interaction concepts of the two design variants. While the LevelTable moves the result set as a whole into the next grade of detail, the GranularityTable gives the user the possibility to explore each result`s set details individually. While both start with an overview of all search results and end with a document browser, the GranularityTable can be adjusted to different grades of detail on one screen.

3 Testing GranularityTable vs. LevelTable

In October 2002 the html-mockups were tested by several users (n=8). Among individual differences in working and searching habits, the majority stated to have very high expectations concerning the work with metadata and that they expect higher efficiency with a working metadata browser.
After the pre-test questionnaire and a video introduction, the users were handed out a script with test tasks to work on. All tests were documented by minute taking, a video camera and a screen-recording.
The overall reception of the SuperTable/Scatterplot framework was good to very good. Some interactions of the Level- and GranularityTable were surprising to the users, though appreciated. Our analysis of the post-test showed, that users were in favor of the GranularityTable. Contradictory to this the results of the test sessions showed that users had more problems to fully understand the interaction concepts of the GranularityTable than of the LevelTable. We argue that, once understood, the joy-of-use is higher with the GranularityTable. Additionally it is possible that users were influenced by the clearer and more aesthetic design of the GranularityTable, and therefore gave it better ratings than the LevelTable. Our test design ruled out learning or last-item remembering effects. Nevertheless this bias in the test design has to be eradicated for the next user tests.
In parallel to the lab tests we started a web-based evaluation. Questions regarded individual search behavior, a virtual search with the two design variants, how the users would interact with them and what they would like to have differently. The participants were asked to download two short introductory videos, and several screenshots correlating to tasks. The sequence of Level-/GranularityTable was randomized to exclude learning and last item remembering effects. 35 users completed the questionnaire, out of them 31 were put into the final evaluation.
Although screenshots are even more limited than the prototypes of the lab evaluation, some results from the former evaluation were confirmed. Throughout the test the effectiveness (measured in correct answers regarding interaction) was higher with the LevelTable than the GranularityTable. A lack of connection between the table visualizations and the Scatterplot was frequently criticized as well.
An interesting result came from the analysis of search behavior and preferences in design. With five separate questions concerning typical search tasks, we wanted to characterise the users in more analytical or more browsing search strategy types (Marchionini 1995). As could be expected, a mixture between both strategies dominated the sample. Only eight users had very clear preferences, five of them we categorized with “only browsing strategy”, three of them with “only analytical strategy”. Interestingly enough, four of the first category absolutely preferred the GranularityTable, and all three of the second category preferred the LevelTable.
We assume that at least for the first steps of an iterative search process the LevelTable can be efficient in analyzing the result set as a whole, maybe find patterns or reformulate/discard the query due to unsatisfactory results. Content is not the primary goal, but filtering and reduction of

the result set is. If then the results are narrowed down to potentially interesting documents, the GranularityTable with its browsing comfort can be used. Now content is the primary goal, modalities can be changed frequently. In this manner, our initially developed scenarios were partly validated by empirical results, though our scenario characters begin the information retrieval process with analytical, very formal and sophistically formulated queries only, while during the iterative retrieval process they become more informal and data driven. Although the evidence should not be weighted too strong, we took it as a hint to handle both design variants equitable.
Using both tables integrated in one search might speed up and ease the visual information retrieval process. This and further interaction concepts are part of the ongoing redesign of VisMeB.

4 Outlook

Throughout the implementation of VisMeB, wide ranges of usability engineering efforts have been taken. With iterative expert evaluations, user tests, questionnaires and user scenarios, a certain level of security in usability questions is reached. The html-mockups were replaced by a java prototypical application, which is constantly monitored and evaluated.
The Scatterplot has now a zoom function, coupled with the result tables. Zoomed documents are placed on top of the tables, the other documents are faded out. Magic Lense Filters in the Scatterplot, inspired mainly by (Fishkin & Stone 1995), can be configured and combined with boolean operators to filter out unwanted documents. Additionally a 3D-Scatterplot is implemented, where data is displayed as transparent cubes. Rotation provides illumination from all directions, a zoom function as well as different selection mechanisms complete the equipment of the 3D-Scatterplot.

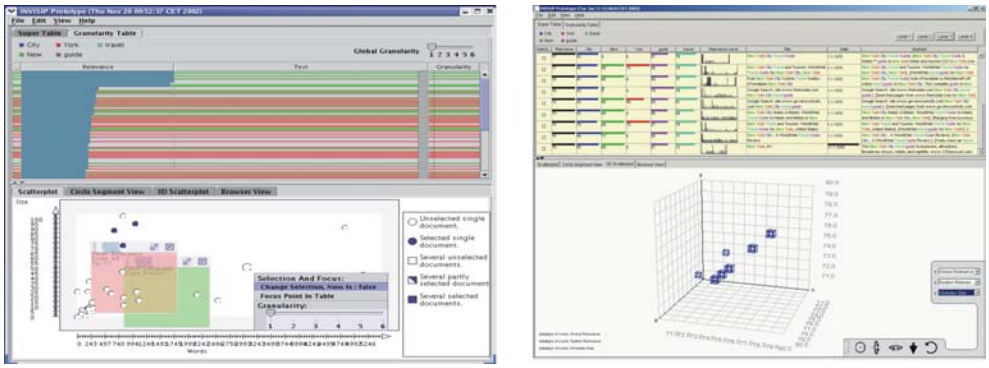

Figure 2: GranularityTable with Magic Lense Filters, Level Table with 3D-Scatterplot

Both result tables now give immediate feedback to user actions. Through a mouse-over-effect, navigational help is given, which is especially useful in both starting levels with low detail. The height of the row is changed, and the user can see this document's details from the next level. In all other levels, moving the mouse over a document highlights the corresponding data point in the Scatterplot and vice versa. Current implementation scenarios for VisMeB regard the use as a geometadata browser, a browser for a movie database and a web search. It will be enhanced by a tool for customization of the visualizations, influenced by ideas from (, Plaisant, & Shneiderman 2000). Users will be able to choose, in which level of detail they want which metadata visualized with which visualization (e.g. Barcharts, RelevanceCurves).
The next steps will be user tests on the java prototype and a comparative evaluation between a form filling and a visual query preview interface. The 2D and 3D-Scatterplots will be compared concerning their effectiveness and efficiency.

References

Ahlberg, C. & Shneiderman, B. “Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, In B. Adelson, S. Dumais, J.S. Olson (Eds.), *CHI 1994: Conference Proceedings Human Factors in Computing Systems,* Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 313-317.

Andrews, K., Gütl, C., Moser, J., Sabol, V., Lackner, W.: “Search Result Visualization with xFIND”.*UIDIS ’01: Proceedings of the Second International Workshop on User Interfaces to Data Intensive Systems*, IEEE 2001

Eibl, M. “DEViD: a media design and software ergonomics integrating visualization for document retrieval” *Information Visualization,* 1, 2002, p.139-157

Fishkin, K. & Stone, M.C. “Enhanced Dynamic Queries via Movable Filters.” In I.R. Katz; R.L. Marks et al (Eds.): *CHI 1995: Conference Proceedings Human Factors in Computing Systems.* Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 23-29

Hearst, M. “TileBars: Visualization of Term Distribution Information in Full Text Information Access.” In I. Katz; R.L. Mack; L. Marks et al. (Eds.): *CHI 1995: Conference Proceedings Human Factors in Computing Systems*. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 59-66.

Klein, P., Müller, F., Reiterer, H. & Eibl, M. “Visual Information Retrieval with the SuperTable + Scatterplot”, *6th International Conference INFORMATION VISUALISATION,* 10-12 July 2002, London.

Marchionini, G. “Information seeking in electronic environments”, *Cambridge Series on Human-Computer Interaction,* Cambridge University Press 1995

North, Christopher L.; Shneiderman, Ben: “Snap-Together Visualizations: Can Users Construct and Operate Coordinated Views*”. International Journal of Human-Computer Studies,* Academic Press, 53(5), November 2000, p. 715-739

Rao, R., Card, S.K. “The Table Lens. Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information.” *in* B. Adelson;S. Dumais; J.S. Olson (Eds.): *CHI 1994: Conference Proceed-ings Human Factors in Computing Systems*. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 318-322.

Rosson, M.B., Carroll, J.M. “Usability Engineering. Scenario-Based Development of Human-Computer Interaction” San Diego (Academic Press) 2002

Veerasamy, A. , Navathe, S. B “Querying, Navigating and Visualizing a Digital Library Catalog” *Proceedings of the Second Annual Conference on the Theory and Practice of Digital Libraries*, 1995