Interactive Data Visualization for Product Search

Mandy Keck

Technische Universität Dresden, Germany
mandy.keck@tu-dresden.de

Abstract. In complex search scenarios such as planning a vacation or finding a suitable gift for a friend, at the beginning the user usually does not know exactly what he is looking for. However, this is the question that most search interfaces present as first step. This research aims to analyze approaches for supporting the user in expressing a search query based on vague motives and ideas and in evaluating the search results in order to find a suitable search result. Various visualization techniques and prototypes are developed to support different stages of the search process and lead to a construction kit for visual search interfaces.

Keywords: Information Visualization, Product Search, Explorative Search, Search Strategies, Faceted Search, Search by Example, Construction Kit

1 Introduction

Complex search tasks such as finding a suitable car, planning a vacation, or identifying the perfect investment opportunity can last days or weeks and usually the user does not know exactly what he is looking for at the beginning. Search engines offer access to large data volumes and various possibilities to interpret the users query like providing corrections and suggestions. Besides these technical advantages, the search paradigm itself did not change a lot during the last years. Most of the conventional web search interfaces use the well-known search box or search forms to express the search query and require the user to transform a possibly vague information need into a specific search query [5]. Users with low experience in the current search domain or with a vague information need have problems to formulate their vague ideas into a concrete query. Besides that, typical search interfaces offer one-dimensional lists with simple sorting and filtering functions [6]. In contrast, the research area of Information Visualization provides various techniques to visualize multidimensional data sets to enhance the quick comprehension, comparison, and analysis of large result sets.

Furthermore, the introduced scenarios correspond to more exploratory forms of search, which require much more diverse strategies, rather than simply submitting a query and seeing a list of matching results [18]. Widely used tools support information access, such as searching on the web, in digital libraries or product databases, but other stages of the information journey are poorly supported at the present [1]. This leads to the need of analyzing and understanding the search process, placing the users and their search behaviors in the focus of the research and develop a richer repertoire of interface solutions to support different stages of the search process.

The aim of this research is to analyze the search process from the users perspective focusing on the use case of a product search. The use cases focus on complex search tasks with a vague information need such as planning a holiday or finding a suitable investment opportunity. Furthermore, different visualization techniques will be evaluated with the aim to enhance the quick interpretation and analysis of the products that are structured as multidimensional data sets, and the quick evaluation and comparison of the results.

The previous considerations lead to the following objectives of my PhD work:

1. the examination of approaches to support the user to express the search query
2. the investigation of different search behaviors and search models to support different search strategies and stages during the search process
3. an in-depth analysis of methods to present the result set to the user that supports a quick interpretation and comparison

The paper is organized as follows: Section 2 addresses background information and state of the art concerning Information Search and Information Visualization. Section 3 outlines the research methodology and section 4 describes a proposed approach to support a product search with vague information need. Section 5 presents briefly different prototypes and evaluation results, which are a basis for a construction kit for visual search interfaces that is introduced in section 6. Finally, section 7 concludes the paper and outlines future work.

2 Background and State of the Art

This research spans several areas, such as Information Retrieval, Information Seeking, Human-Computer Interaction and Information Visualization. Whereas Information Retrieval focuses on the technologies that support the finding and presentation of information, Information Seeking is primarily concerned with the seeking of information and focuses on the users and their search activities [1]. Latter will be introduced in section 2.1, whereas section 2.2 focuses visualization techniques for multidimensional datasets in the research area of Information Visualization.

2.1 Information Seeking

Search activities can be distinguished in Lookup and Exploratory Search [14]. Lookup describes the most basic kind of search tasks such as fact retrieval, known item search, and question answering. Exploratory Search is a more complex

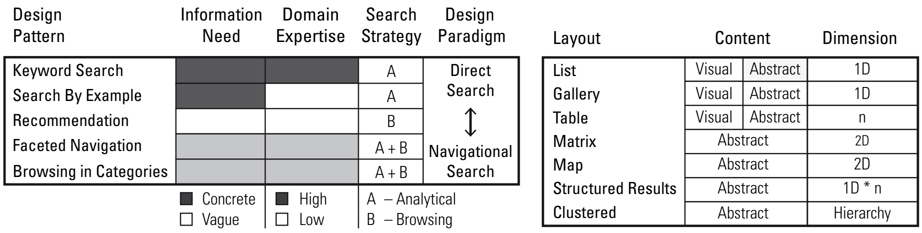

Fig. 1. Design Patterns (left) and Layout Patterns (right) in Product Search

process that usually starts with a vague information need and therefore requires multiple iterations of learning, investigation, and reformulation of the search query. The use case of a product search corresponds to exploratory search, with the difference that product search is not an open-ended search but aims to conclude with finding a suitable product. Exploratory search scenarios often start with a vague information need and usually blend two search strategies: analytical and browsing [14]. In contrast to the formal, analytical strategies - that depend on careful planning and iterative query reformulation - browsing strategies are more informal and interactive, can foster serendipity and depend on recognizing relevant information [6].

Furthermore, two search paradigms are well established in the web search world: Direct Search and Navigational Search [17]. Direct Search allows users to simply write their queries in a text box and became enormously popular with web search engines, such as Google and Yahoo! Search. Text boxes and search forms are well-suited for lookup-scenarios, in which the user has a concrete idea of the desired product (e.g. looking for a flight to London). In contrast, Navigational search systems provide guidance through the use of a taxonomy [17].

I analyzed different design patterns in the context of product search and assigned to the introduced search strategies and paradigms (see Fig. 1, left). Design Patterns have emerged as recurring solutions to common problems and can be adapted to the current context [15]. Most reviewed e-commerce website provide the well-known keyword search paradigm with design patterns such as *Autosuggest*, *Autocomplete*, *Autocorrect*, *Instant Results*, *Partial Matches*, *Search Within*, *Scoped Search* and *Advanced Search* (cp. [15], [16]). These design patterns address searches for users with a concrete information need and a better domain expertise and are not in the scope of this research. In contrast, the design patterns *Search By Example*, *Recommendation*, *Faceted Navigation* and *Browsing in Categories* provide alternatives to the keyword paradigm and can be used for users with a vague information need and little domain knowledge and are more suitable for further investigation in context of this research.

2.2 Information Visualization

Furthermore, I analyzed different e-commerce websites to identify patterns to present the result sets. Products were either presented as an image, when the visual appearance is important e.g. in case of travel destinations or clothes, or they were presented in an abstract way, mostly through textual descriptions or using simple diagrams like bar charts or line graphs. Thus, not many attributes of a product can be presented and one-dimensional lists and galleries are still the mostly used patterns (see Fig. 1, right). In contrast, the research area of Information Visualization offers various techniques for visualizing multidimensional data, which are needed in the context of product search. Keim distinguishes between geometric techniques (e.g. scatterplots, parallel coordinates), icon-based techniques (e.g. star plots, chernoff faces), and pixel-oriented techniques, where each data value is mapped to a colored pixel [12]. Using icon-based techniques, each data record becomes a small independent visual object and data attributes are mapped to graphical attributes of each glyph, such as size, shape, color and orientation [3].

3 Research Methodology

The research methodology is divided into 5 main steps:

1. Analysis of the current state of the art in Information Seeking to understand the search process and its strategies and to identify limitations and desiderata
2. Analysis of the current state of art in Information Visualization with respect to the visualization of products with multidimensional attributes
3. Proposal of a model to support the product search with a vague information need and the aim to find a suitable product
4. Design, prototyping, and evaluation of different search approaches targeting different data sets, user experiences, and information needs
5. Decomposition of the prototypes and assignment to different contexts in product search with the aim to develop a construction kit for visual search interfaces that is providing patterns, which suit to different data structures and targets and is supporting the designer in giving inspiration for the development of new interfaces

4 Motive-based Search

To get a better understanding of how people seek information and to describe the process of information search from the users perspective, Kuhlthau performed a series of studies and identified distinct phases and emotions unique to each phase [13]. Particularly in the initial stages, uncertainty and anxiety are an integral part of the process, followed by feelings of confusion, doubt, and frustration in the exploration phase. Although the first phases include the most complex tasks for the user, most search applications invest most of their effort in the later phases [16] and the user is forced to express his vague ideas and motives as a specific query, which the system can understand.

The goal of this research is to investigate strategies and methods to support the first stages of the process of information search. Therefore, it is concentrated on complex search tasks, such as planning a vacation, in which the user is unsure of what he is looking for at the beginning. Motive-based Search refines Explorative Search by specifying the motive and the aim of the process to find a suitable product. A motive can be defined as the reason for a search as well as particular conditions such as how much a product should cost and is usually influenced by emotions and interests of the user. Because this motive is the starting point of the search task, this type of task is called *Motive-based Search* in this thesis.

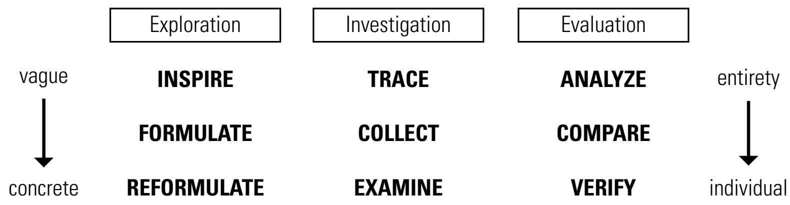

Fig. 2. Aspects and Tasks of Motive-based Search

Based on a workshop with potential users and the research of Marchioni and Russel-Rose, who consider search as a holistic process, integrating findability with analysis and sensemaking [14] [16], three phases of the motive-based search could be identified: **Exploration**, **Investigation** and **Evaluation** that contains different tasks that should be supported during the search process and are explained by using the scenario of planning a vacation (see Fig. 2) [9].

In the beginning of the search, the user decides to book a vacation without a concrete idea where he wants to go. Based on his motive for the vacation, some basic conditions and constraints have to be met. In this example, he is looking for a cheap short trip in the near future and all travels by air can be quickly discarded because of his fear of flying. In the phase of Exploration, the user is unsure at the beginning of his search and needs inspiration and guidance to start the search process. To support the user in creating new ideas, the interface has to provide functions that broaden his scope of information (see *Inspire* in Fig. 2). During the search, the search idea is getting more concrete and leads to *Formulation* and *Reformulation* tasks.

When finding some interesting topic, e.g. a trip to Paris, Rome or Madrid, the users task is to read thematically relevant information, and to relate this information with previous knowledge in order to extend the personal understanding of the topic. The interface can support the user by offering functions to construct and organize his knowledge space (see *Investigation* in Fig. 2). The tasks that should be supported in this phase are described as follows: *Trace* (shows the user, where he has been already and helps to orientate in the information space), *Collect* (supports the user with collecting and organizing his findings), *Examine* (gives more information about an unknown finding and supports the learning process).

Finally, the user has to grasp the possibilities of combining bits of information and different alternatives. The task of the interface in the last phase is to help the user to narrow the scope and to create a focus (see *Evaluation* in Fig. 2). The Evaluation is necessary to judge the value of an item or item collection with respect to the search goal. It is supported by tasks that are leading from *Analyzing* the entire result set and *Comparing* two ore more items to identify similarities and differences; to a *Verification*, that is used to confirm if one item meets all specific criteria.

5 Visual Search Interfaces for Product Search

Ten prototypes were developed, which focus on different search scenarios and stages of the motive-based search process. In this section, four of these prototypes are presented based on different data sets and visualization techniques[1]. Each concept supports a particular search strategy to find the desired product in a very large database with structured or semi-structured data. Further on, different search domains such as image search, travel search, and financial products are considered. Since the domains entail different user requirements and expectations, the resulting interface concepts support emotional or analytical decisions.

5.1 Relation-based Concept

The first concept is based on a folksonomy as data structure organized in a multi-dimensional classification scheme. The developed DelViz (Deep exploration and lookup of Visualizations) concept supports different search tasks such as finding suitable visualizations for a given context, and analysis of the underlying structured data set to discover relationships between the search results [8]. To support these search tasks, the application offers two flexible areas: the representation of the clasffication scheme on the left-hand side, and the visualization projects presented as thumbnails on the right-hand side. A splitter in the middle can be used to expand one of these two areas to change the level of detail on either side (see Fig. 3, top).

5.2 Facet-based Concept

The second concept focuses on data sets that are structured with a faceted classification. It is based on two visualization techniques that allow the representation of multidimensional data across a set of parallel axes: parallel coordinates [7] and

[1] http://www.visual-search.org provides the associated videos and prototypes

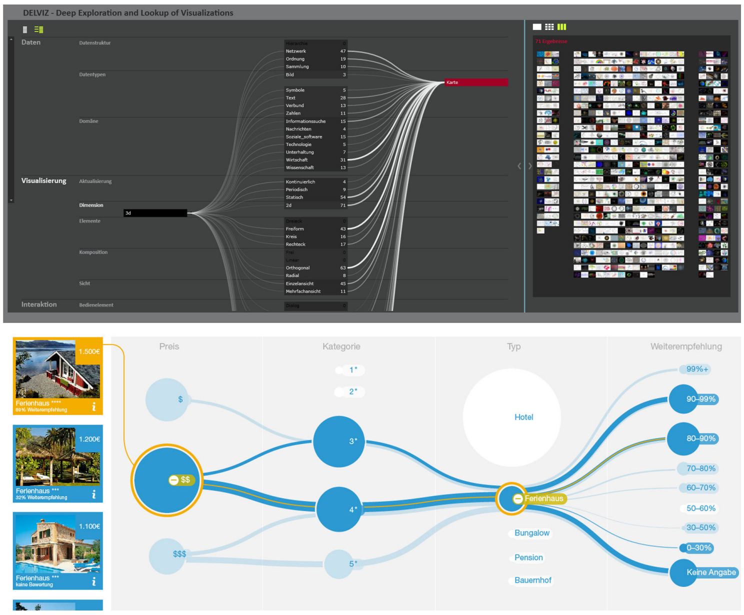

Fig. 3. Top: DelViz Prototype (top): tags can be selected (red) or removed (black) on the left-hand side, the right-hand side represents generated subsets, Bottom: Parallel Sets for Travel Search

parallel sets [2]. Figure 3 (bottom) shows an interface designed for travel search based on the principle of parallel sets. The interface concept combines principles of Faceted Browsing with the visualization method of parallel sets to support additional analytical tasks. We developed the interface with the parallel set and parallel coordinates technique. A user test compared both variations with each other and has shown that parallel sets are faster and easier to understand with regard to faceted searches and analysis tasks, whereas parallel coordinates have advantages in comparison tasks and similarity-based searches [10].

5.3 Recommendation-based Concept

We designed a recommendation-based concept to support and trigger emotionally driven decisions and uses an ontology of concepts as data set that describe a holiday, such as warm, beach, party and culture. Fig. 4 (top) shows the interfaces getInspired for a travel search with a vague information need. Instead of a direct query on the attribute of the result set, the user communicates his preferences to the system through a selection of concepts represented by expressive images. Based on the previous decision of the user, the system decides which concepts are presented in the next refinement step. A user study with 29 participants and 12 different search tasks was conducted to compare the recommendation-based approach to a strict navigational concept selection. The study measured time and clicks to solve a given search task and indicated that the recommendation-based approach was faster and more efficient than the navigational concept. Furthermore, the results show an obvious trend of increased user experience as well as inspiration while the search result quality has not dropped [4].

5.4 Example-based Concept

Fig. 4 (bottom) shows a visual approach for a similarity search in a financial product scenario and is created for users with low domain expertise, who cannot express their information need with filters, like the first and second concept. The

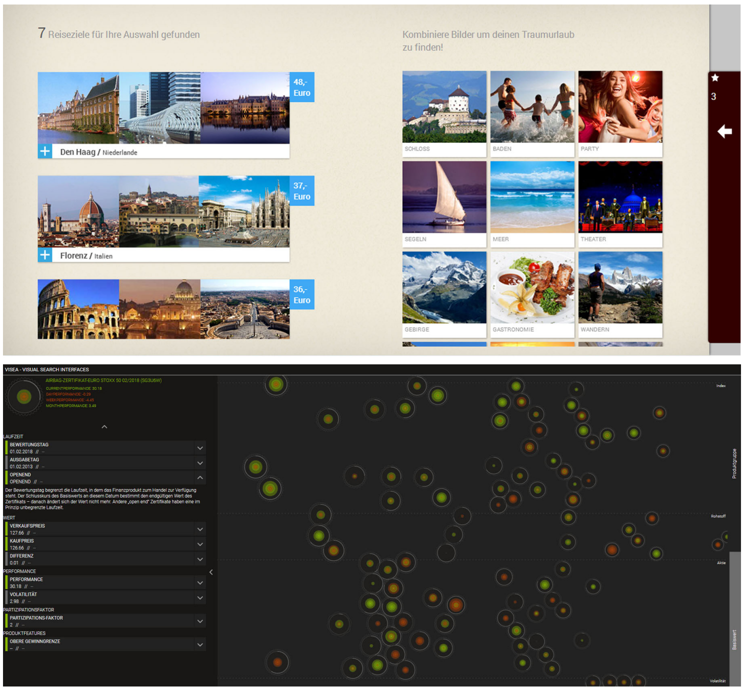

Fig. 4. Top: getInspired Interface with visual concepts for travel searches, Bottom: Visual Similarity Search: attributes of the reference product can be selected or deselected, which influence the glyphs on the right side

user starts the search with a given example of one financial product on the left-hand side and gets more information about its properties, such as risk, maturity, and other special features for this single product. Then the user can decide for each attribute if this criterion is important and can select and deselect them, which influences the similarity algorithm. On the right-hand side, each circular glyph represents a single product and depicts its most important features, such as performance over time and investment term.

An early prototypical implementation suggests that this kind of interface is well suited for the paradigm of search by example and the browsing of a large structured or semi-structured dataset. Whilst the overall design of the interface may remain the same for any kind of structured dataset, the iconic representation should be adapted to the use case [11].

6 Construction Kit for Visual Search Interfaces

Based on the previous introduced prototypes a construction kit was developed, which aims to support the search process with the aid of visualization. The construction kit contains different *building blocks*, introduced in Figure 5. These building blocks can be combined to a *pattern*, that is subdivided into three parts:

What: describes the data input
Why: describes the task that has to be solved
How: describes how the pattern is designed

The elements of the parts "What" and "How" are influenced by the visualization taxonomy used in the multidimensional classification scheme for information visualizations introduced in [8] and the elements of "Why" refer to the tasks identified in section 4.

These patterns can be composed to *construction plans*, that enables to describe complex search interfaces with the help of three different connectors:

Successive: the patterns are shown successively
Juxtaposed: the patterns are shown next to each other
Superimposed: the patterns are combined with each other in one view

Furthermore, it is possible to combine individual building blocks with *reference blocks* to indicate the influence on another pattern (see Figure 6, reference blocks "B" and "C").

In order to explain this composition, Figure 6 shows a construction plan that has been reverse engineered from a search interface using parallel sets (see section 5.2). The underlying data set is a *faceted classification*. The **Search Task** is to *formulate* search queries and to *analyze* how many results are left in each node. Pattern A is visualized by using the building blocks *area*, a *rectangular* grid and *bars* to compare the sizes of data sets. *Highlighting* is used to refer to connected streams, and the bars function as a filter to *reduce* the result set. The last interaction method influences pattern B and C, hence, it is combined with two *reference blocks* (shown as small rectangles attached to the bottom right of a pattern. Pattern B deals with the underlying faceted classification as well. Facets can be *nominal* (location), *ordinal* (ranking) or *quantitative* (price). The **Search Task** is to *analyze* the correlations between these facets and to support the *reformulation* of search queries. It uses a *parallel plot* as *Layout Structure* with *rectangular Grid* but with *flows* instead of lines to represent a set of results. Single flows can be *highlighted*. It is combined with pattern A by a *superimposed* connector, hence they overlap in one visualization. The bars serve as axes for

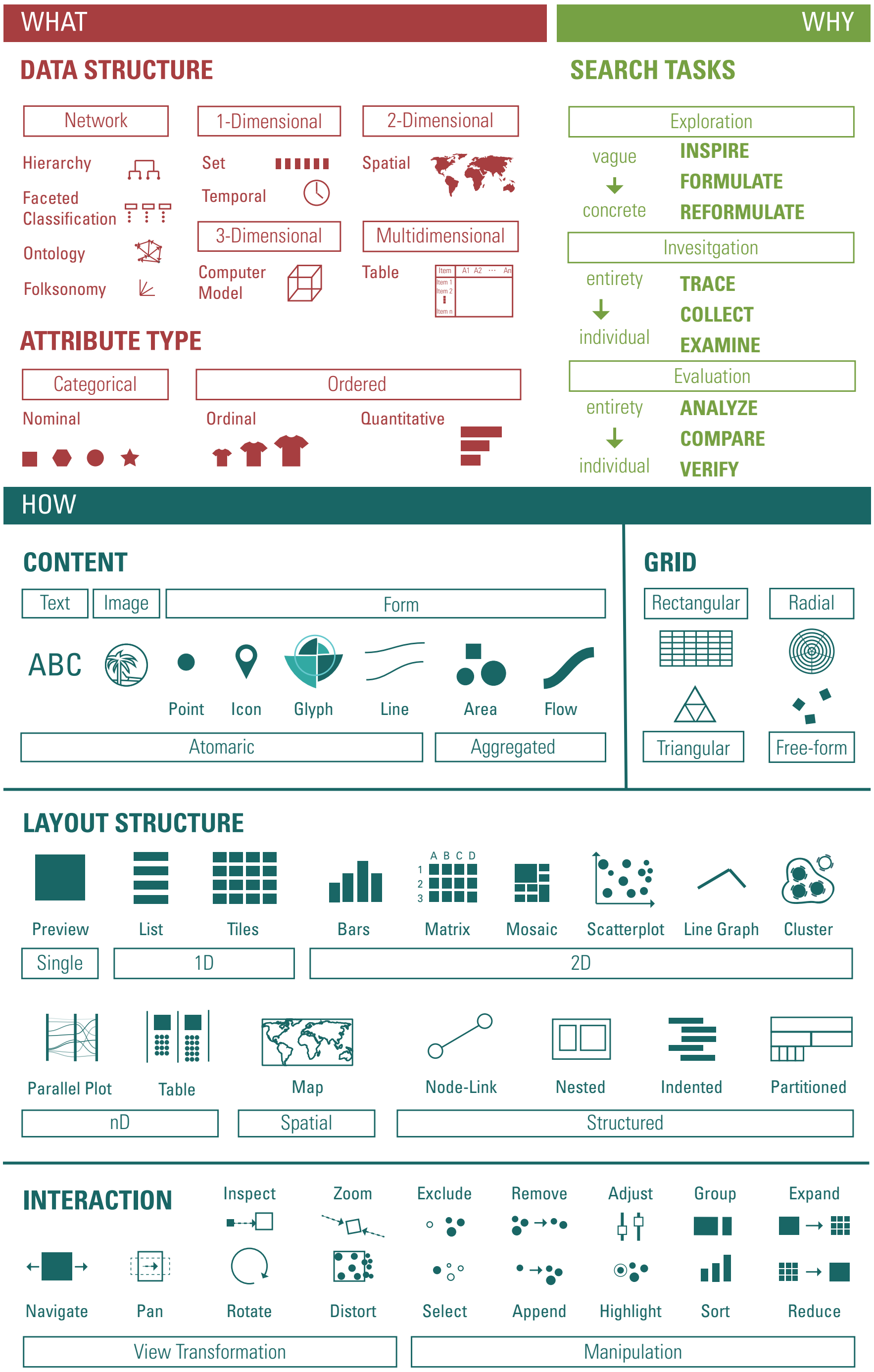

Fig. 5. Building Blocks of the Construction Kit

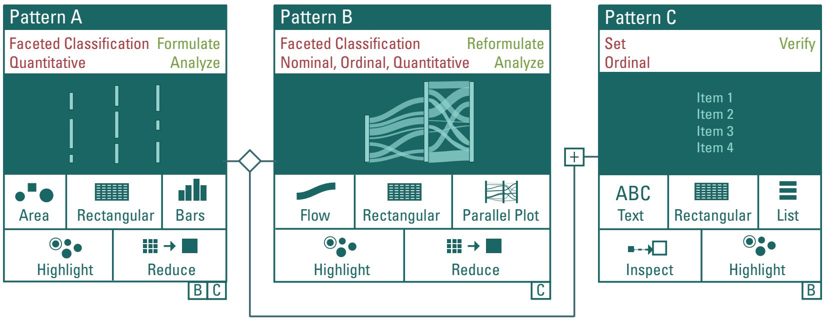

Fig. 6. Construction Plan of the Parallel Set Prototype

the parallel plot layout and also as filter in pattern B, which is indicated by the *reference block* attached to pattern A, referencing pattern B. Also in pattern B, the interaction method *reduce* combined with the *reference block* indicates the influence of the filter on pattern C. The underlying **Data Structure** of pattern C is an ordered (*ordinal*) result *set* with the **Search Task** to *verify* individual items. All items are ordered in a *list* of a title (*text*) in a *rectangular* grid. More details are shown for each item by using the *inspect* interaction method. The pattern is combined with the *juxtaposed* connector to pattern A and B and is influenced by their filtering. The interaction *select* in pattern C also influences pattern B, by selecting single items in the list that are highlighted in the other visualization.

7 Conlusions and Outlook

This paper presents the current stage of the PhD thesis. The presented construction kit, presents the last phase of the PhD and serves the purpose to support the designer in quickly creating new visual search interfaces by giving him or her a set of elements, which can be easily combined with each other. Resulting patterns can be used for reuse and adaptation in different contexts. The patterns are conceptually simple and provide a solid foundation for reuse and redesign. Furthermore, they can be networked in a very flexible way to create complex search interfaces. However, the construction kit itself provides only syntactic support for this combination process by offering the elements and the rules to combine the building blocks. The semantic level is concerned with the content and its meaning, which will be addressed by providing example patterns based on existing solutions. This will be addressed in the last phase of my PhD work.

Acknowledgments. I would like to thank my supervisor, Prof. Rainer Groh for the support and guidance provided in this research. Parts of this research has been supported by the European Union and the Free State Saxony through the European Regional Development Fund (ERDF).

References

1. Blandford A. and Attfield, S.: Interacting with Information. Morgan and Claypool Publishers (2010)
2. Bendix F., Kosara R., Hauser H.: Parallel Sets: Interactive Exploration and Visual Analysis of Categorical Data, Transactions on Visualisation and Computer Graphics Vol.1 (2006)
3. Borgo, R., Kehrer, J., Chung, D., Maguire, E., Laramee, R., Hause, H., et al.: Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. Eurographics State of the Art Reports. Eurographics Association. (2013)
4. Both, A., Nguyen, V., Keck, M., Herrmann, M., Kammer, D., Groh, R. & Henkens, D.: Motive-based Search using a Recommendation-driven Visual Divide and Conquer Approach. In: IADIS International Journal on WWW/Internet (2014)
5. Doerk, M., Williamson, C. and Carpendale, S.: Navigating tomorrow's web: From searching and browsing to visual exploration. ACM Trans. Web 6, 3, Article 13, 28 pages. New York, USA (2012)
6. Hearst, M.A.: Search User Interfaces. New York: Cambridge University Press. (2009)
7. Inselberg, A. & Dimsdale, B.: Parallel coordinates: A tool for visualizing multi-dimensional geometry. In Proc. of IEEE Visualization, 361 378 (1990)
8. Keck, M., Kammer, D., Iwan, R., Taranko, S., Groh, R.: DelViz: Exploration of Tagged Information Visualizations Conference Informatik 2011 - Interaktion und Visualisierung im Daten-Web, Berlin (2011)
9. Keck, M., Herrmann, M., Both, A., Gaertner, R., Groh, R.: Improving Motive-Based Search. Conference Proceedings of the First International Conference on Distributed, Ambient, and Pervasive Interactions - Volume 8028, Springer-Verlag New York, Inc., New York, NY, USA, 2013, (2013)
10. Keck, M., Herrmann, M., Both, A., Henkens, D.. Groh, R.: Exploring Similarity, Proceedings of the 16th International Conference on Human Interface and the Management of Information. Information and Knowledge in Applications and Services, 8522 , Springer-Verlag New York, Inc., New York, NY, USA (2014)
11. Keck, M., Herrmann, M., Henkens, D., Taranko, S., Nguyen, V., Funke, F., Schattenberg, S., Both, A., Groh, R.: Visual Innovations for Product Search Interfaces. 1st International Workshop on Future Search Engines at INFORMATIK 2014 (2014)
12. Keim, D.: Designing Pixel-Oriented Visualization Techniques: Theory and Applications. IEEE Transactions on Visualization and Computer Graphics 6, (2000)
13. Kuhlthau, C. C.: Inside the Search Process: Information Seeking from the Users Perspective. Journal of the American Society for Information Science and Technology, 42(5), 361371. (1991)
14. Marchionini, G.: Exploratory Search: From Finding to Understanding. Communications of the ACM, 49(4):4146, (2006)
15. Morville, P. and Callender, J.: Search Patterns. O'Reilly. Canada (2010)
16. Russell-Rose, T. and Tate, T: Designing the Search Experience: The Information Architecture of Discovery, Morgan Kaufmann (2012)
17. Tunkelang, D.: Dynamic category sets: An approach for faceted search. ACM SIGIR. 6, (2006)
18. Wilson, M. L., Kules, B., Schraefel, M.C., and Shneiderman, B.: From Keyword Search to Exploration: Designing Future Search Interfaces for the Web. Found. Trends Web Sci. 2, 1 (January 2010), 1-97.(2012)