

# 10471 Abstracts Collection

## Scalable Visual Analytics

— Dagstuhl Seminar —

Daniel A. Keim<sup>1</sup> and Stefan Wrobel<sup>2</sup>

<sup>1</sup> Universität Konstanz, DE  
keim@uni-konstanz.de

<sup>2</sup> Fraunhofer IAIS - St. Augustin, DE  
Stefan.Wrobel@iais.fraunhofer.de

**Abstract.** From 21.11. to 26.11.2010, the Dagstuhl Seminar 10471 “Scalable Visual Analytics” was held in Schloss Dagstuhl – Leibniz Center for Informatics. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

**Keywords.** Visual Analytics, Visualization, Data Analysis, Discovery Science, Information Visualization

### 10471 Executive Summary – Scalable Visual Analytics

Today’s fast growing data sets and information availability pose new challenges for data processing. Visual Analytics addresses the problem by combining automatic data analysis techniques with interactive visualizations and thus, integrating the user where necessary effectively in the process.

The Scalable Visual Analytics seminar was a fertile meeting in which researchers from diverse backgrounds met. It included industry and academia, senior and junior researchers, multi-national representation, and people coming from several disciplines. The diversity resulted in interesting and useful discussions, which will help to shape the future of the versatile research area of Visual Analytics.

The seminar included multiple presentations and discussions which helped to exchange domain knowledge and steer future research activities. Besides, several working groups during the seminar not only identified future research directions in the field of scalable visual analytics but also initiated new joint projects. In total, plans for three position papers, two overview papers to outreach to other communities, and three EU FET Open Projects were drafted. Furthermore, three workshops as satellites of conferences that cover specific application areas were planned to further disseminate the work and provide a platform for ongoing discussions and activities.

*Keywords:* Visual Analytics, Visualization, Data Analysis, Discovery Science, Information Visualization

*Joint work of:* Keim, Daniel A.; Wrobel, Stefan

*Extended Abstract:* <http://drops.dagstuhl.de/opus/volltexte/2011/2939>

## **Exhaustive visual search for information in multi-dimensional data sets**

*Georgia Albuquerque (TU Braunschweig, DE)*

Goal of this research project is to develop and evaluate a fundamentally new approach to exhaustively search for, and interactively characterize any non-random mutual relationship between attribute dimensions in general data sets. To be able to systematically consider all possible attribute combinations, we propose to apply image analysis to visualization results in order to automatically pre-select only those attribute combinations featuring non-random relationships. To characterize the found information and to build mathematical descriptions, we rely on interactive visual inspection and visualization-assisted interactive information modeling. This way, we intend to discover and explicitly characterize all information implicitly represented in unbiased sets of multi-dimensional data points

*Joint work of:* Magnor, Marcus; Theisel, Holger; Albuquerque, Georgia; Lehmann, Dirk J.

*Full Paper:*

<http://graphics.tu-bs.de/projects/spp1335/index.html>

## **ViAMoD: Visual Spatiotemporal Pattern Analysis of Movement and Event Data**

*Gennady Andrienko (Fraunhofer IAIS - St. Augustin, DE)*

Movement data are difficult to analyze due to their amount, irregular sampling, missing values, large diversity of types of movement, and the need to consider movement in spatial and temporal context. Within the ViAMoD project, we develop theoretical foundations for the analysis of movement data, address various types of movement data, develop methods for joint analysis of movement data and context, and find approaches to overcome the scalability limitations.

After identifying real world data sets and problems, we developed elements of the theoretical framework and proposed several methods for analyzing movement and event data.

*Keywords:* Visual analytics, movement, geo, time

*Joint work of:* ViAMoD project consortium

## Role of Mashups, Cloud computing, Parallelism for Visual Analytics

*Loretta Auvil (Univ. of Illinois - Urbana-Champaign, US)*

In this talk, I promote the use of three technologies that will benefit visual analytics: mashups, cloud computing, and parallelism. First, mashups will encourage the breaking down of silos with data driven, service oriented solutions. Infrastructure tools like component based frameworks can be used to integrate tools and techniques from many different teams. There are many of these frameworks available, but I focus on the one that we developed. In Meandre, components can be created by leveraging services and APIs from tools. This component based strategy enables mashups to be created for analytics and visualizations by other application areas more easily. Secondly, cloud computing approaches can provide cycles for on-demand computation and/or delivery of tools and services. In Meandre, we have created components that can access our cloud system to launch computation. Thirdly, parallelism is necessary to speed up our large scale data crunching needs. In Meandre, we have implemented MapReduce so that a component can be automatically parallelized to improve performance.

*Keywords:* Mashup, cloud computing, parallelism, meandre

## Interactive Motion Segmentation

*Benjamin Berkels (Universität Bonn, DE)*

Interactive motion segmentation is an important task for scene understanding and analysis. Despite recent progress state-of-the-art approaches still have difficulties in adapting to the diversity of spatially varying motion fields. Due to strong, spatial variations of the motion field, objects are often divided into several parts. At the same time, different objects exhibiting similar motion fields often cannot be distinguished correctly.

Here, we propose to use spatially varying affine motion model parameter distributions combined with minimal guidance via user drawn scribbles. Hence, adaptation to motion pattern variations and capturing subtle differences between similar regions is feasible.

The idea is embedded in a variational minimization problem, which is solved by means of recently proposed convex relaxation techniques. For two regions (i.e. object and background) we obtain globally optimal results for this formulation. For more than two regions the results deviate within very small bounds of about 2 to 4 % from the optimal solution in our experiments. To demonstrate the benefit of using both model parameters and spatially variant distributions, we show results for challenging synthetic and real-world motion fields.

*Joint work of:* Nieuwenhuis, Claudia; Berkels, Benjamin; Rumpf, Martin; Cremers, Daniel

*Full Paper:*

<http://www.springerlink.com/content/v4835ppu1h283108/>

## **Visualization of Subspace Clustering Results**

*Enrico Bertini (Universität Konstanz, DE)*

In this work we introduce the problems of visualizing the output of subspace clustering algorithms. The main problem of this algorithms stems from the fact that each cluster has an associated set of records and dimensions thus making their visualization not trivial. We describe ClustNails a system that addresses these needs and permit the users to asses the quality and meaning of subspace clustering results. We also present a preliminary systematization of alternative models for tight human-machine integration in the construction of reliable clustering models. We propose two paradigms: visualization of alternatives and visually-aided algorithm steering.

*Keywords:* Subspace clustering, visualization, visual analytics, human-computer interaction

## **Visualization of and interaction with complex graphs on large-scale an high-resolution displays: models, metaphors, and interaction paradigms**

*Sven Boettger (TU Kaiserslautern, DE)*

With software growing in complexity safety analysis for software and embedded systems becomes increasingly difficult. We present a new sunburst layout approach for component fault trees and forests. This enables the developer to get a better understanding about the failure propagation from basic faults in the system to top events or unwanted system faults. Additionally we use metrics to manipulate the layout making important parts more visible than unimportant parts of the graph structure and support common cause analysis.

*Keywords:* Safety Analysis, Component Fault Trees, Sunburst Layout, Failure Propagation

## **Designing Applications for Interactive Surfaces**

*Susanne Boll (Universität Oldenburg, DE)*

Large scale and touch sensitive tabletops and surfaces offer promising opportunities for Visual Analytics applications.

They enable users to directly interact with visualized data by touch and facilitate delving into the data without using abstracting input devices such as mice and keyboards. Nevertheless, the development of corresponding applications for interactive surfaces is still a major challenge, as common development process models do not consider the specific characteristics of such devices sufficiently. As a solution, we have developed SCiVA, an iterative, user centric design process that is specifically tailored to the specific characteristics of interactive tabletops and surfaces, which will be presented in this talk.

*Keywords:* Surface computer, visual analytics, design methods, human centered design

*Joint work of:* Boll, Susanne; Hesselmann, Tobias

## Slides for the Infrastructure for VA

*Jean-Daniel Fekete (INRIA Saclay - Orsay, FR)*

VA applications are currently created from scratch, re-implementing algorithms and visualization components. This presentation describes a path towards solving this problem.

*Keywords:* Software architecture

## Presentation slides: The Grand Challenge of Scalable (Cognitive) Visual Analytics

*Brian D. Fisher (Simon Fraser Univ. - Surrey, CA)*

A discussion of possible roles for cognitive science in visual analytics. I conclude that existing cognitive science disciplines offer substantial support for visual analytics research but that support can be increased by the creation of cognitive science methods that focus more precisely on visually enabled reasoning. This poses a challenge for both cognitive science and visual analytics to find ways to work together. I suggest that this can only take place if a collaboration infrastructure of conferences and possibly journals are created to support that collaboration.

*Keywords:* Cognitive science, visually-enabled reasoning, interdisciplinary collaboration, translational research

## Visual Analytics as a Translational Cognitive Science

*Brian D. Fisher (Simon Fraser Univ. - Surrey, CA)*

Visual analytics is a new interdisciplinary field of study that calls for a more structured scientific approach to understanding the effects of interaction with complex graphical displays on human cognitive processes.

Its primary goal is to support the design and evaluation of graphical information systems that better support cognitive processes in areas as diverse as scientific research and emergency management. The methodologies that make up this new field are as yet ill-defined. This paper proposes a pathway for development of visual analytics as a translational cognitive science that bridges fundamental research in human/computer cognitive systems and design and evaluation of information systems in situ. Achieving this goal will require the development of enhanced field methods for conceptual decomposition of human/computer cognitive systems that maps onto laboratory studies, and improved methods for conducting laboratory investigations that might better map onto real-world cognitive processes in technology-rich environments.

Please email me at [bfisher@sfu.ca](mailto:bfisher@sfu.ca) for a preprint, the journal publisher will not accept Creative Commons license

*Keywords:* Visualization; human-computer interaction; visual analytics; information systems; translational research; cognitive systems; visually-enabled reasoning; augmented cognition

*Joint work of:* Fisher, Brian D.; Arias-Hernández, R.; Green, T.M.

## Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics

*Brian D. Fisher (Simon Fraser Univ. - Surrey, CA)*

Studying how humans interact with abstract, visual representations of massive amounts of data provides knowledge about how cognition works in visual analytics. This knowledge provides guidelines for cognitive-aware design and evaluation of visual analytic tools. Different methods have been used to capture and conceptualize these processes including protocol analysis, experiments, cognitive task analysis, and field studies. In this article, we introduce Pair Analytics: a method for capturing reasoning processes in visual analytics. We claim that Pair Analytics offers two advantages with respect to other methods: (1) a more natural way of making explicit and capturing reasoning processes and (2) an approach to capture social and cognitive processes used to conduct collaborative analysis in real-life settings. We support and illustrate these claims with a pilot study of three phenomena in collaborative visual analytics: coordination of attention, cognitive workload, and navigation of analysis.

*Keywords:* Pair analytics, qualitative methods

*Joint work of:* Arias-Hernández, R.; Kaastra, L.; Green, T.M.; Fisher, Brian D.

*Full Paper:* <http://drops.dagstuhl.de/opus/volltexte/2011/2938>

*Full Paper:* <http://www.computer.org/portal/web/csdl/proceedings/h#2>

## Visually guided parameter choice in selected analysis tasks

*Joachim Giesen (Universität Jena, DE)*

We consider three examples for parameterized data analysis methods, namely (a) linear support vector machines specialized to conjoint analysis, (b) singular value decomposition, and (c) alpha shapes for point cloud data in Euclidean space. The parameter in (a) is a regularization parameter that trades off a regularization and a data term (a good value for this parameter prevents over- and underfitting), the parameter in (b) is the number of singular vectors, and the parameter in (c) is the radius (alpha) of balls centered at the data points. We demonstrate interactive tools that aid the user in choosing good parameter values for all three examples.

*Keywords:* Conjoint analysis, singular value decomposition, geometric multi-scale analysis

## Visual Analysis of Biomedical Documents

*Carsten Goerg (UCHSC- Aurora, US)*

Finding relevant publications in the large and rapidly growing body of biomedical literature is challenging. Search queries on PubMed often return thousands of publications and it can be a tedious task to filter out irrelevant publications and choose a manageable set to read. We have developed a visual analytics system, named Bio-Jigsaw, which acts like a visual index on a document collection and supports biologists in investigating and understanding connections between biological entities. We apply natural language processing techniques to identify biological entities such as genes and pathways and visualize connections among them via multiple representations. Connections are based on co-occurrence in abstracts and also are drawn from ontologies or annotations in digital libraries. We demonstrate how Bio-Jigsaw can be used to analyze a PubMed search query on a gene related to breast cancer resulting in over 1500 primary papers.

*Keywords:* Visual analytics, investigative analysis, entity identification, language processing, biomedical literature

*Joint work of:* Goerg, Carsten; Tipney, Hannah; Verspoor, Karin; Baumgartner, William; Cohen, Kevin; Stasko, John; Hunter, Lawrence

## Modularity and Infrastructure in the eHumanities

*Gerhard Heyer (Universität Leipzig, DE)*

By looking at an example from the BMBF funded project eAQUA - a project that applies advanced text mining to digitized classical Greek and Latin texts - we propose that searching for the rare cases can be considered a research paradigm in the Humanities to which visual analytics can make valuable contributions.

The talk further explores how this approach can be modularized and made available in a distributed research infrastructure.

*Keywords:* EHumanities, modularization, infrastructure, text mining

*Full Paper:*

<http://www.eaqua.net/en/index.php>

*See also:* <http://www.eaqua.net/en/public.php>

## Scalable Visual Analytics for Video Data

*Markus Hoeferlin (Universität Stuttgart, DE)*

The major goal of the project "Scalable Visual Analytics for Video Data" is to develop a system that supports users in the interactive analysis of video data. In this talk, we discuss the principal structure of our proposed system and present the project results of the past two years according to this structure. Afterwards, we discuss current work in progress, future directions, past and recent collaborations, and possible generalizations of the project.

## Visualization of Wine Tasting Notes

*Andreas Kerren (Linnaeus University - Växjö, SE)*

This talk is concerned with the visualization of wine tasting notes in order to support linguistic analyses. The ultimate aim of our collaboration partners is to explore the descriptions used for human sensory perceptions. The ISOVIS group at Linnaeus University developed a prototype implementation that gives a better insight into the language used in wine tasting notes. A number of concepts for text analyses are discussed in this talk, and several approaches for text visualization and statistical information are presented.

*Keywords:* Information Visualization, Word Trees, Scatter Plots, Wine, Perception

## SPP project presentation "Scalable Visual Patent Analysis"

*Steffen Koch (Universität Stuttgart, DE)*

Patent retrieval is a very important task for a variety of interest groups. Research in this field aims to increase recall and to reduce the time needed for search and analysis. We propose a visual analytics approach to achieve these goals.



So far, access to a variety of backend systems has been set up with patent data acquired from the European Patent Office. These backend systems include search indexes for textual information, as well as a relational database for patent metadata. Both, NLP/retrieval methods and visual techniques have been successfully integrated to improve upon retrieval quality in a visual patent retrieval interface. Current work in this project focuses on (semi)automatic (multilingual) query expansion for patent retrieval, while future aspects of our work will also take into account classification tasks including the user-in-the-loop, e.g. by applying active learning.

*Keywords:* Visual analytics, patent retrieval, patent analysis, dokument retrieval

*Joint work of:* Koch, Steffen; Jochim, Charles; Ertl, Thomas; Schütze, Hinrich

## Continuation of the VisMaster Initiative

*Joern Kohlhammer (FhG IGD - Darmstadt, DE)*

This presentation reflects on the VisMaster initiative and possible ways to continue the initiative through FET Open proposals, a European consortium structure, and other funding opportunities.

FET Open proposal ideas that I "distilled" from the VisMaster Research Roadmap included:

- Semantics-based Methods for the Management and Visual Analysis of Incomplete, Inconsistent, and Erroneous Data
- Ambiguity-Tolerant Visual Analytics Approaches to Combine Data from Conflicting Sources
- Integrated Processing and Analysis of Diverse Spatio-Temporal Data at Multiple Scales
- New Evaluation Methodologies Integrating Research on Design, Perception, and Visualisation

*Keywords:* VisMaster, Visual Analytics, Funding Opportunities, FP7, FET Open

## Network-based interactive navigation and analysis of large biological datasets

*Hans-Peter Lenhof (Universität des Saarlandes, DE)*

Since all kinds of biochemical processes can be modeled as graphs or networks, the analysis and visualization of biological networks plays a central role in the life sciences and especially in the field of systems biology.

Most applications here are characterized by the fact that heterogeneous large-scale data sets have to be integrated, analyzed, and visualized. Hence, a framework that facilitates and accelerates the development and implementation of visual analysis tools for such applications is of the utmost importance. In this talk, we discuss the components, features, and functionality of our framework for visual analytics of large-scale biological data in a network context based on a selected application from cancer research. Key application here is the interactive identification of the abnormal biological processes occurring in tumor cells, the deregulated pathways. To this end, we discuss the requirements for a suitable visual analytics framework to answer this question, starting with the data management concepts and functionality for the integration, storage, and retrieval of data sets. Then we present two novel approaches for detecting deregulated paths and subnetworks. While the first approach for paths, a dynamic programming algorithm, is based on Gene Set Enrichment Analysis, the second approach for finding maximally deregulated subgraphs is a branch-&-cut approach based on an integer linear programming formulation of the problem. Finally, we demonstrate the functionality of our visualization tool BiNA with a special emphasis on biologically motivated layouts and interactive components required for the study of deregulated paths and subgraphs. Based on this framework and visualization tool we obtained biologically interesting results on different cancer-related datasets, illustrating the application potential of the methods and tools developed here.

*Keywords:* Visual analytics, bioinformatics, cancer, deregulated pathways, BiNA

*Joint work of:* Lenhof, Hans-Peter; Kaufmann, Michael; Kohlbacher, Oliver; Weikum, Gerhard

## Understanding large, complex software systems through visual analytics

*Peter Liggesmeyer (TU Kaiserslautern, DE)*

The evaluation of software characteristics based on models or measurements is an established research area. Its importance increases due to growing software size, critical software properties (e.g., safety, security) and new application domains (e.g., ubiquitous computing / ambient computing), that are characterized by complex behavior (e.g. dynamic aspects).

Understanding characteristics of such large, complex software systems is increasingly difficult both for the developers and for additional persons, that might be interested, e.g. users or consultants in certification authorities. Software is an immaterial product that extracts itself from an intuitive evaluation. Software solutions occur in a variety of applications domains, that require to determine certain software properties.

Safety-critical embedded systems need to be certified by certification authorities, before being released for public use. This requires a technique that is capable to display detailed information on safety properties of a large software system. At the same time, this technique has to avoid generating an information overload.

Systems from the ubiquitous computing / ambient computing domain are likely to be extraordinary complex with respect to dynamic aspects. Furthermore, security issues might be a major concern.

It is not sufficient to be able to determine software properties. In addition, an appropriate representation of the data is required, in order to make sure, that information is not lost or misinterpreted. It is difficult, to keep the overview while perceiving all the detailed information at the same time.

The goals of the research project are

- to develop an appropriate visualization concept, that is applicable to large, complex software systems,
- to investigate representations (metaphors) for immaterial software properties (e.g. safety, security, dynamic behavior),
- to build a prototype system, in order to be able to validate the research results.

It is necessary to visualize the software characteristics in such a way, that it is possible to change dynamically from overview representations to detailed information. Relevant characteristics on the detailed level must remain visible to a certain extent also in the summary view. An appropriate approach might be to use techniques, that are used in spacial information systems. These are characterized by similar problems. Data on the detailed level (e.g. population densities) are to be exchanged dynamically in real time by geographical data at the overview level. It should be investigated, whether these procedures are applicable to the problem domain addressed here. Furthermore, suitable metaphors are to be developed as means for the representation, since natural metaphors for safety, security, dynamic aspects of ubiquitous computing systems, etc. do not exist.

At the beginning, the research will be focused on visualization of safety and reliability properties. These are extraordinary important in embedded systems development. Since techniques are available, that can be used to determine these properties, the research can be focused on their visualization. Furthermore, standards are available, that define certain procedures for safety evaluation. These might help to develop the solutions.

The results will be expanded and transferred to additional properties and application domains. It might be, for instance, interesting to apply it to ambient computing systems or to visualize properties of software development processes.

*Keywords:* Software Engineering, Visualization, Metaphor, Complexity

*Joint work of:* Liggesmeyer, Peter; Ebert, Achim

## SCALABLE Visual Analytics ... Everything Breaks

*Richard May (Pacific Northwest National Lab., US)*

We, the visual analytics community, claim to deal with large and complex information spaces. We in fact do this in many ways at least to some degree. The challenge we face will be two fold. The idea of ŠlargeŠ is growing rapidly with peta and exa scale data flows entering into organizational vocabulary. The major challenge is that as we develop visual analytics systems for large and complex data sets everything will become fragile. Scaling across data that ranges in orders of magnitude is a significant issue for algorithms, visualization, interaction, hardware, analytical methodologies, and even human reasoning.

## A Matter of Time and Interactions: Interactively Exploring Time-Oriented Data

*Silvia Miksch (TU Wien, AT)*

Time is an important data dimension with distinct characteristics that is common across many application domains. This demands specialized methods in order to support proper analysis and visualization to explore trends, patterns, and relationships in different kinds of time-oriented data interactively. A variety of different visualization and analysis methods for time-oriented data have been developed. However, how do these methods account the special characteristics of time and interactions? How are the complex structures of time tackled in visual representations, analysis, and the visualization process? Which interaction methods are provided to explore time-oriented data? And how do these methods scale?

*Keywords:* Time-Oriented Data, Interaction Techniques, Visual Analytics

*Joint work of:* Miksch, Silvia; (Wolfgang Aigner, Heidrun Schumann, Christian Tominski)

## VAExpress: Visual Analytics for large and heterogeneous life science data with emphasis on expression data

*Kay Nieselt (Universität Tübingen, DE)*

Current high-throughput technologies allow to assess transcriptomes in unprecedented detail. Microarrays are most commonly used to monitor the expression of known genes in order to understand their function in a global context. The resulting data is highly heterogeneous and large-scale. When analysing expression data automatic analysis as well as visual exploration play an important role. The aim of VAExpress is to develop a scalable visual analytic framework for

integrative analysis of these diverse expression data. For this purpose we have started to develop SpRay. Central visualization is an adaptation and extension of parallel coordinates. SpRay offers a visual analysis of the original data as well as automatically derived statistical data in one common visual analytics data space. When visualizing large-scale high-dimensional data in 2 dimensions, interesting patterns can be subject to been overdrawn. To address the question of overdrawing we propose several different approaches. The first makes use of motion as a mapping technique which in our opinion, has not yet to been sufficiently exploited in the field of multidimensional visualization. Three techniques, flickering, weaving and mimic motion, are currently exploited and applied to expression data. The second topic is more closely related to the interactive analysis of expression data as one of the main issues of our project. We propose to use a semisupervised clustering algorithm as the computational basis and to combine this with parallel coordinates to evaluate the cluster result as well as to define and also assess constraints to be sent to the underlying cluster algorithm. The major advantage is that the influence of the constraints can immediately be seen on the screen and allows for an immediate evaluation. Thirdly we evaluate different dimension reduction techniques, linear as well as non-linear ones to map high-dimensional data two two or three dimensions.

*Keywords:* Expression data, Bioinformatics, Parallel Coordinates, Clustering, Motion, Boost Clustering, Dimension reduction

*Joint work of:* Nieselt, Kay; Pritzkau, Albert; Polatkan, Aydin Can; Scheuermann, Gerik

## State and Future Prospects of Visual Readability Analysis

*Daniela Oelke (Universität Konstanz, DE)*

A common challenge when producing a text, is to write it in a way that it is easy to read and understand by the target community. This includes aspects like ensuring contextual coherency, avoiding unknown vocabulary, difficult grammatical structures, or misspellings.

In this talk I introduce our VisRA tool for visual readability analysis. The tool was designed in a way that the user does not only see which passages are difficult to read but can also investigate why this is the case. In future work we would like to enhance the tool with additional features such as measuring writing style, consistency, or even analyze whether the argumentation is stringent. One goal is thereby to find out if the concept of approximating semantic text properties with low-level features does also work for more complex aspects. Furthermore, we plan to investigate if there is a correlation between readability scores and review scores.

*Keywords:* Visual analysis, document analysis, text analysis, readability

## PerceptionCognition

*Margit Pohl (TU Wien, AT)*

The design of Visual Analytics systems representing large quantities of information is a challenging problem. Information should be organized on the screen in such a way that it supports human perceptual and reasoning processes. Cognitive theories provide guidance how to achieve this. Human perception is an active process of hypothesis formulation and visual queries. Therefore, interaction with the system should be provided. The representation of time-oriented data is used as an example to illustrate relevant issues in this area. Either time is represented by space ("timelines") or by time itself (animation). Both solutions offer different modes of interacting with the system. They both have advantages and disadvantages from a cognitive point of view. Approaches from cognitive psychology like visuospatial thinking can provide important input to provide design decisions for systems representing time-oriented data.

*Keywords:* Perception cognition

## Visual Analytics and Emotion (ad hoc talk)

*Margit Pohl (TU Wien, AT)*

Emotion plays an important role in interaction processes with computers. In HCI, research concerning user experience tries to clarify issues related to emotion and aesthetics. In information visualisation, research indicates that users interact longer with aesthetically pleasing visualisations (see e.g. Cawthon and Vande Moere 2007). On the other hand, some users indicate that beautiful visualisations might be misleading and convey information not conforming to reality.

*Keywords:* User experience, aesthetics

## Efficient Visual Analysis of Dynamic Medical Image Data

*Bernhard Preim (Universität Magdeburg, DE)*

Dynamic medical image data allow assessing metabolism and other functional aspects. They bear a great potential for diagnosis. The target user group (radiologists) is very well trained but the data quality is low due to many artifacts, e.g. from breathing, heart beat, low spatial resolution, and noise. We focus on perfusion data where the contrast agent distribution in the body is measurement with MRI or ultrasound. The diagnosis task is to characterize breast and brain tumor with respect to malignancy. We developed and refined a visual analytics pipeline including registration (motion correction), segmentation, decomposition of tumors, data analysis of perfusion parameters, and visual exploration,

e.g. with glyphs. The data analysis is a feedback system between trained expert knowledge and evaluation. Extensive statistical analysis of >130 lesions in >70 patients revealed that heterogeneity of tumors better helps to characterize them compared to any individual perfusion parameter.

*Keywords:* Glyph rendering, registration, segmentation, perfusion imaging, visual analytics, medical imaging

*Joint work of:* Preim, Bernhard; Toennies, Klaus; Glaßer, Sylvia; Schäfer, Sebastian

## Choosing the right parameter

*Waqar Saleem (Universität Jena, DE)*

A parametrized machine learning method's "solution path" is the range of values its parameter is allowed to take on. Choosing the right parameter value is both important and non-trivial. We present two of our software tools to visually assist an analyst in this task in different contexts. The first tool visualizes similarities among entities in a corpus as the parameter used in the similarity computation changes. Similarity is computed asynchronously using Singular Value Decomposition (SVD) and is parametrized by the number of Singular Vectors. The second tool explores the solution path of a Support Vector Machine (SVM) based method to analyze conjoint data. Different properties of the solution are visualized along the path of the regularization parameter of the underlying SVM. Finally, we present a method to scale shape retrieval which incorporates insights we gained from our work on the above two tools.

*Keywords:* Conjoint analysis, solution path, parameter selection, SVD, SVM, similarity

## Topology-based Visual Analysis of Information Spaces

*Gerik Scheuermann (Universität Leipzig, DE)*

Visualization of large document collections has to cope with a very large amount of term vectors in a document cloud. On the NLP side of the joint project, we have worked on reducing this amount by introducing the notion of a term's context volatility as a new measure for computing the "highly discussed" topics over a period of time. On the visualization part, we have elaborated on the notion of topological visualization in order to plot topics and their related documents in a document landscape.

*Joint work of:* Scheuermann, Gerik; Heyer, Gerhard

## Project Report: Visual Feature Space Analysis

*Tobias Schreck (TU Darmstadt, DE)*

In this report, we detail the motivation, recent results, work-in-progress, and future work of the part project "Visual Feature Space Analysis". The motivation for our work is the availability of multiple, often complementary descriptor spaces, for implementing similarity search and analysis for many complex data types. Our idea is to make use of multiple descriptor spaces in Visual Analytics applications. We research interactive approaches, possibly supported by comparative views, and semi-automatic approaches. We aim to support retrieval and cluster analysis tasks, on different complex data types. Recent results include a multi-descriptor visual analysis system for graph data; approaches for comparative analysis of alternative projection and cluster views; and a visual cluster analysis system for spatiotemporal data.

*Keywords:* Visual Analytics, high-dimensional data, descriptors, cluster analysis, similarity search

*Joint work of:* Schreck, Tobias; Bremm, Sebastian; Von Landesberger, Tatiana

## Zoomable Cell

*Michael Schroeder (TU Dresden, DE)*

The zoomable cell integrates microscopy images and protein structure and interaction data to facilitate seamless navigation across many scales. The talk highlights two aspects of the project: First, the construction of multi subunit complexes; Second, biomedical image search and navigation.

Regarding the former, we developed an algorithm which exploits protein interactions to constrain the explosion in combinatorial docking. Regarding the latter, we built a library of 13 Mio images and filtered to 745000 relevant images. The images are indexed with an ontology which allows filtering. Additionally, we have investigated various measures of image similarity and implemented an interactive system to navigate along multiple access of image similarity and to built mini classifiers to find images suitable to a users choices.

*Keywords:* Bioinformatics, zoom, morph, image search, ontology, textmining, 3D strucutre