Knowledge-assisted EHR visualization for cohorts

Paolo Federico, Albert Amor-Amorós, and Silvia Miksch

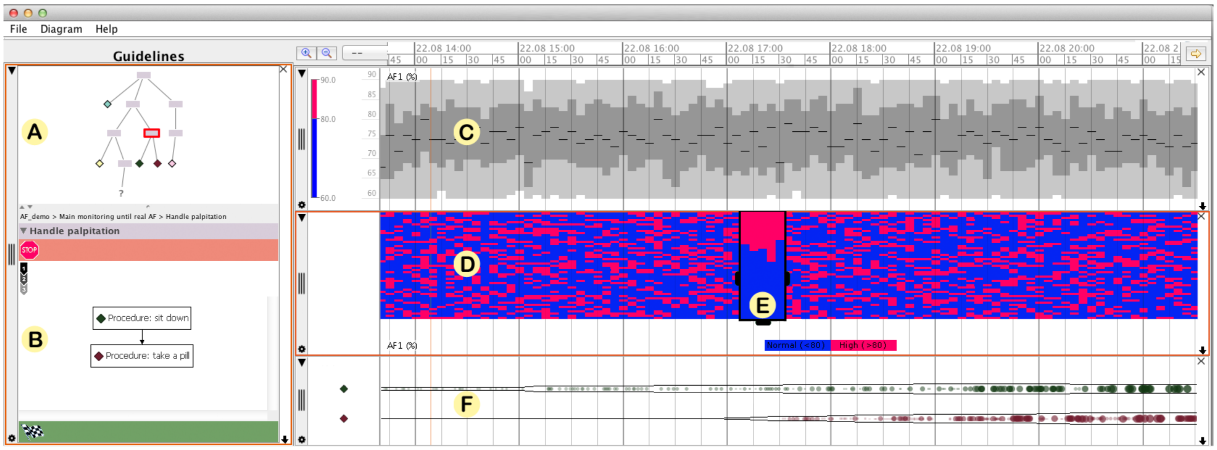

Fig. 1. The GUI of *Gnaeus*, a guideline-knowledge-assisted EHR visualization for cohorts. (A) The hierarchical structure of the clinical guideline comprising subplans and actions is shown as a tree visualization with a layered top-down layout. (B) The procedural knowledge of the selected subplan is shown as a node-link hierarchical task network. (C) The raw data of that parameter, which is relevant for the selected subplan, is aggregated over the patients' cohort and shown as a streaming box-plot. (D) The data is also abstracted according to the declarative knowledge of the subplan and visualized as lifelines. (E) An interactive grouping lens reconfigures the abstractions to visualize their distribution as vertical barcharts. (F) The execution of clinical actions and their compliance to the guideline recommendations are shown in an aggregated visualization.

Abstract—The advanced visualization of electronic health records (EHRs), supporting a scalable analysis from single patients to cohorts, intertwining patients' conditions data with executed treatments data, and handling the complexity of time-oriented data, is an open challenge of visual analytics for health care. An integrated approach addressing and meeting the challenge would enable a better analysis of the EHR data and a deeper comprehension of the health care process, thus providing several benefits in terms of reduction of costs and risks to the patient, and improved quality of care. According to the knowledge-assisted visualization paradigm, we propose a solution that leverages the domain knowledge acquired by clinical experts and formalized into computer-interpretable guidelines (CIGs), in order to improve the automated analysis, the visualization, and the interactive exploration of EHR data of patients cohorts. The declarative and procedural knowledge constituting the CIGs is used for automatically checking the compliance of the treatment to the medical recommendations, abstracting and visualising patients' data with respect to the intentions of the treatment plans, and filtering relevant data for specific analytical tasks; in this way, the analyst can get insights about the clinical history of multiple patients and assess the effectiveness of their health care treatments.

Index Terms—Electronic Health Records, Computer-interpretable guidelines, Knowledge-assisted visualization, Visual Analytics

1 Introduction

In recent years, the diffusion of Electronic Health Records (EHRs) has been growing, partially due to specific public health policies and legislative interventions, such as the National Programme for IT in the United Kingdom since 2002, the Health Information Technology for Economic and Clinical Health Act of 2009 in the United States, and the directive 24/2011/EU for cross-border healthcare in the European Union. Besides facilitating the online data transfer amongst hospitals, clinics, and care providers during the treatment, the increasing adoption of EHR systems has made available large amounts of data about patients' conditions and their care pathways. Retrospective analysis of this data can be exploited to assess the effectiveness of treatments and identify complex patterns and special cases, in order to improve the overall quality of healthcare. Several interactive information visualization techniques and systems have been proposed to visually explore EHR data, gain insights, and form and validate hypotheses. Generally, the effective utilization of these systems requires analysts to rely on their domain knowledge, in order to interpret raw data and deduce the overall health status of a patient, as well as to assess the administered treatments and compare them with evidence-based best practices.

In this context, we propose a solution for the visualization of EHR data of patients' cohorts. Our main contributions are:

- a knowledge-assisted visual analytics approach, that leverages the knowledge condensed into computer-interpretable clinical guidelines in order to drive analysis, visualization, and interaction;
- a proof-of-concept implementation, that demonstrates the advantages of such approach.

- *Paolo Federico, Albert Amor-Amorós, and Silvia Miksch are with the Vienna University of Technology. E-mail: {federico,amor,miksch}@ifs.tuwien.ac.at.*

Proc. of IEEE VIS 2014 Workshop on Visualization of Electronic Health Records.

2 Related Work

Rind et al. [14] conducted an extensive and systematic survey about interactive information visualization approaches to EHR exploration and querying; they also identify future research directions, like designing smoother transitions between the analysis of single patients and the analysis of patients' cohorts in order to support the comparison of special cases versus general trends.

Evidence-based clinical practice guidelines (CPGs) are sets of statements and recommendations used to improve health care by providing a trustworthy comparison between treatment options in terms of risks and benefits according to patient's status. They condense in a standardized narrative form the complex domain knowledge underneath the clinical practice. Their formalization as computer-interpretable guidelines (CIGs) enables the implementation of guideline-execution engines and decision-support systems, assisting professional care providers during the daily practice [10]. Several visualization techniques for CIGs have been proposed in the literature; they are generally aimed to visually support their acquisition and specification (e.g., AsbruView [8]), or to enable direct visual editing (e.g. Gesher [17]).

Systems like CareVis [2] and CareCruiser [7] intertwine the visualization of CIGs with executed treatment and patient's health status, in order to assess the effects of the former onto the latter. Bodesinksy et al. [4] present a system for the visual analysis of compliance, i.e. the extent to which an executed treatment fulfills the recommendations a certain patient is eligible for [13]; their approach is inspired by the Visual Analytics paradigm, i.e. the tight integration of automated analysis, visualization, and user interaction [18].

Despite the wealth of visualization systems and techniques present in the literature, there are still many open challenges of visual analytics for healtcare [1]. In this work we focus particularly onto three of these challanges: the simultaneous exploration of both single patients and cohorts data, the intertwined analysis of patients' conditions and treatment data, and an appropriate support for the time-oriented nature of EHR data.

3 A KNOWLEDGE-ASSISTED VISUALIZATION APPROACH

To address these challenges, we present *Gnaeus*[1], a guideline-knowledge-assisted EHR visualization for cohorts. Building upon the aforementioned systems [4, 7], aimed to enable an integrated visual analysis of EHRs and CIGs, we have designed a solution that makes this integration tighter, and exploits all the domain knowledge condensed into the clinical guideline to assist the visual analytics process. For many complex visual analytics tasks, indeed, users generally rely on their knowledge to interpret data; the analysis of EHR data, in particular, can only be performed by expert users, who have the necessary domain knowledge to interpret patients' data and assess executed treatments.

According to the knowledge-assisted visualization approach [5], *Gnaeus* exploits expert knowledge to better support parts of the visual analytics process. It does not acquire clinical knowledge specifically for the visualization, since knowledge is usually acquired from medical experts into narrative-form CPGs to be used in daily practice, and also formalized into CIGs to be processed by decision-support systems. By placing the CIG at the core of *Gnaeus*, we are able to leverage the domain knowledge to inform the automated analysis, the visualization, and the user interaction for EHR data.

3.1 Clinical guidelines as a knowledge base

Gnaeus has been specifically designed to use guidelines written in Asbru, an intention-based and time-oriented language for CGI knowledge representation [9]. "Intention-based" means that an essential element of an Asbru guideline are the intentions, i.e. the goals expressed at various level of abstraction; intentions can be understood as temporal patterns of actions or states to be achieved (or avoided), and can be temporally annotated by the means of complex time-oriented constructs.

Additional elements of an Asbru guideline are effects, describing the functional dependencies between clinical actions and patients' parameters (e.g., administration of antipyretic decreases the body temperature). Intentions, temporal abstractions, and effects can be seen as the declarative knowledge formalized in the CIGs, as they describe what can be observed and what is to be to achieved.

[1] *Gnaeus* is a Latin forename. Many prominent Romans bore this name, in particular *Gnaeus Petreius*, senior centurion and commander of the first cohort.

An Asbru guideline also comprises the preferences, the conditions, and the plan body. They represent the procedural knowledge of the guideline, recommending how to proceed (activating or aborting subplans, performing specific clinical actions) in order to accomplish the guideline intentions according to patients' parameters.

3.2 Automated analysis

In our design, both the declarative knowledge and the procedural knowledge are exploited to drive the automated analysis. The declarative knowledge, specified as guideline intentions, is exploited to compute temporal abstractions: raw numerical data (bio-signals, measured parameters, other patients' parameters) are abstracted into nominal values according to states, gradients, rates, or patterns consisting of any combination of the previous [15]. The obtained nominal values incorporate the domain knowledge and are thus easier to interpret within the specific context of the guideline than the raw data. A rule-based engine, using the procedural knowledge of plan bodies, preferences, and conditions, processes treatment data and patients' conditions to check the compliance of the executed treatment to the recommendations [4]. The compliance analysis is performed individually for each patient, but the engine computes also aggregate statistics, such as the ratio of compliant/non-compliant occurrences of a specific action across the whole cohort.

3.3 Visualization

Gnaeus adopts coordinated multiple views for visualizing EHR data as well as the procedural knowledge of the CIG. The hierarchical structure of the guideline is visualized as a tree diagram with a top-down layered layout, whose nodes represent subplans and leaves represent clinical actions . The logical structure of a subplan is shown as a node-link diagram of a hierarchical task network. Figure 1 A-B shows the hierarchical structure and the logical structure of a small illustrative guideline for the self-management of atrial fibrillation (AF) with a pill-in-the-pocket approach: in case of palpitation and measured AF probability above the 80% threshold, the patient is recommended to reduce physical activity, take a pill and, if the condition persists, call a doctor.

The procedural knowledge of the guideline can be used to visually analyze the synchronization of executed subplans, aggregated over a cohort. Asbru subplans, indeed, can be characterized by various temporal constraints: *parallel* (to be started together), *sequential* (one after another, in a given order), *any order* (one after another, in any order), and *unordered* (no synchronization constraints). Figure 2 shows a modified tree layout encoding the synchronization constraints, while the edge color represent the average execution time; comparing the expected execution order with the actual execution time, the user can check the synchronization of the executed treatment, identifying temporal patterns and outliers. Since the layout integrating synchronization information is less space efficient than a simple top-down layered layout, it can be toggled on demand.

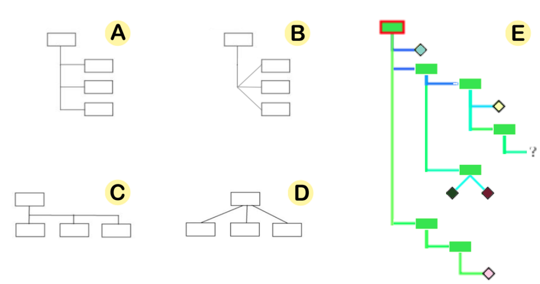

Fig. 2. The layout of the guideline hierarchy tree can reflecr the synchronization of subplans as defined within the guideline: (A) sequential ordered, (B) sequential any order, (C) parallel, (D) unordered. (E) The color shading, according to the average execution time (blue=earlier, green=later), enables the comparison between the expected synchronization and the actual execution.

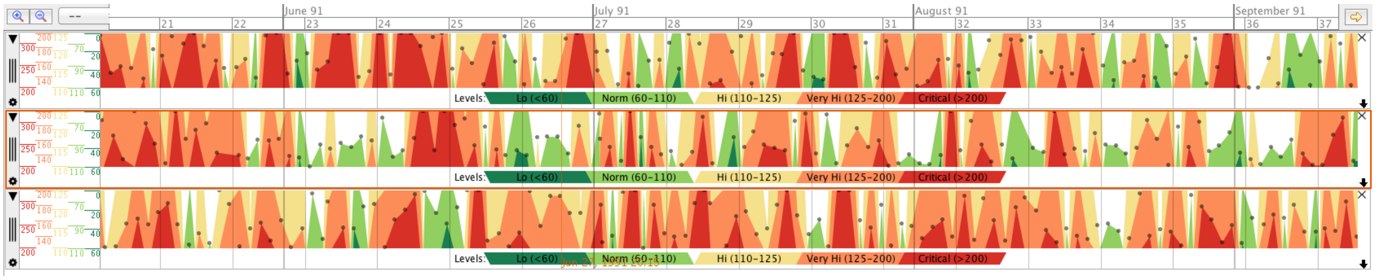

Fig. 3. The glycemia of three patients visualized as qualizon graphs, an integrated visualization of raw data and their knowledge-based abstractions.

The temporal abstractions, reducing a numeric (univariate or even multivariate) parameter into a nominal variable, enable a compact visualization of time-oriented EHR data. For small cohorts, *Gnaeus* supports qualizon graphs, a space-efficient visualization combining quantitative data and qualitative abstractions for single patients (Fig. 3). Qualizon graphs are based on the well-known horizon graphs, but they extend them with non-uniform bands corresponding to the value ranges of state abstractions; they are as fast and accurate as horizon graphs for raw data, but also support the integrated visualization of state abstractions [6].

For more complex abstractions, or for larger cohorts demanding a more compact visualization, *Gnaeus* provides a pixel-based lifelines visualization; Figure 1 D shows the qualitative abstractions of a parameter such as the probability of atrial fibrillation: each horizontal line corresponds to a patient, and it is coloured according to the temporally abstracted states (blue = normal, magenta = high).

The raw data for a numeric parameter can also be aggregated over all the patients of a cohort and visualized as a streaming box-plot (Fig. 1 C); it shows the five-number statistical summary: the mean is mapped to the black line, the 25th and 75th percentile are mapped to the dark gray bars, the minimum and maximum are mapped to the light grey bars.

Executed actions are aggregated over the cohort and visualized as transparent circles along the time-axis (Fig. 1 F): the number of occurrences of an action within an interval is mapped to the alpha channel, while the number of patients to which the action has been administered is mapped to the radius. The height of the bar represents the number of patients within the interval who were eligible for the recommendations: the white space between the action circle and the ends of the eligibility bar represents the number of non-compliant patients (in other words, the number of missing actions or not fulfilled recommendations).

3.4 Interaction

Given the large scale, the multivariate nature, and the temporal complexity of EHR data, specific interaction methods are needed to support user's intentions and enable data exploration. A first set of interaction techniques provided by *Gnaeus* is aimed at facilitating the transition between analysis of single patients and cohorts. A magic lens [19] reconfigures the arrangement of the temporal view of qualitative data. Outside of the lens, each lifeline represents the history of a single patient; within the lens, the lines are grouped by abstraction, thus enabling a quick overview of the distribution of abstractions across the population in terms of bar charts (Fig. 1 E).

Conversely, a fish-eye interaction allows the user to focus on a single patient. When hovering upon the lifeline of a patients in the qualitative view, this line is magnified (Fig. 4 B), the corresponding quantitative data is overlaid on top of the streaming box-plot (Fig. 4 A) and the corresponding treatment data is also highlighted in the context of the cohort(Fig. 4 C). These interaction techniques enable a direct comparison between a given patient and the rest of the cohort.

The system provides also knowledge-assisted interactions, supporting specific tasks in the context of a guideline. Since an EHR can in principle contain a large amount of multivariate time-oriented data for each patient, the guideline can be used as an index to browse the EHR data both across the different variables and along the time axis. Exploiting the plan-parameter dependency specified in the CIG declarative knowledge, when the users selects a subplan in the guideline views (Fig. 1 A-B), only the set of relevant parameters and actions is shown in the temporal views (Fig. 1 C-D-E). Moreover, the time axis can switch from absolute time to relative time, and all patients within the cohort are aligned according to the execution time of the subplan or action selected by the user.

3.5 Implementation

Gnaeus has been designed to be integrated within the wider architecture of the MobiGuide system [12], an intelligent decision-support system for patients with chronic illnesses; in that context, guidelines are stored in and retrieved from the DEGEL digital library [16], mappings between guideline concepts and EHR archetypes are managed by the KDOM knoweldge-data mapper [11], abstractions are computed by the IDAN temporal mediator [3], and compliance analysis is performed by the RoMA reasoner [13].

A stand alone prototypical implementation, with its own analytical components, has been developed for small-scale datasets and simple demonstration scenarios.

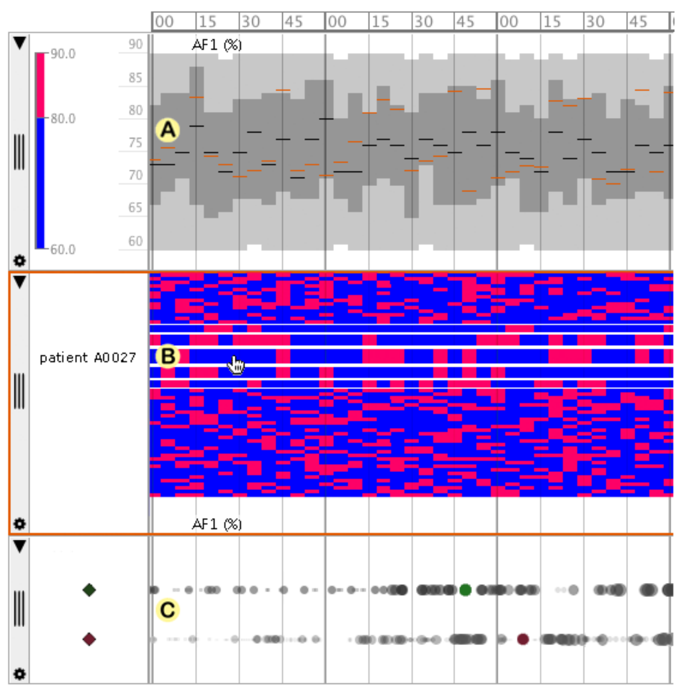

Fig. 4. The temporal views demonstrating user interaction: (A) the raw data of the selected patient are overlaid upon the streaming box-plots as orange lines; (B) a vertical fish-eye distortion enables a closer examination of the abstractions of a single patient's data; (C) the treatment of the selected patient is highlighted while the rest of the cohort is grayed out.

4 Conclusion

We have presented *Gnaeus*, a guideline-knowledge-assisted EHR visualization for cohorts that exploits the domain knowledge of clinical computer-interpretable guidelines to support the visual analytics process and drive automated analysis, interaction, and visualization.

Asbru-formulated clinical guidelines are a rich source of declarative and procedural knowledge; other particular aspects of these guidelines, such as effects, can be further investigated and utilized for assist analysis and visualization. As next step, we also plan to address the computational and visual scalability of *Gnaeus*, in order to support larger cohorts and enable an efficient visual comparison between them. Moreover, the overall approach and specific techniques need to be evaluated by studying real-world use cases and by collecting feedback from domain experts, such as clinical practitioners and researchers.

Acknowledgments

The authors wish to thank their partners in the MobiGuide project for fruitful discussions and inspiration. We also wish to thank Peter Bodesinsky, Michael Fischer, Tobias Froihofer, Alexander Rind, and Jürgen Unger for sharing source-code of their previous works or coding parts of the prototypical implementation of *Gnaeus*.

This work was funded in part by the European Commission (grant number 287811, MobiGuide project).

References

[1] W. Aigner, P. Federico, T. Gschwandtner, S. Miksch, and A. Rind. Challenges of time-oriented data in visual analytics for healthcare. In J. J. Caban and D. Gotz, editors, *Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Healthcare*, page 4. IEEE, IEEE, 2012.

[2] W. Aigner and S. Miksch. Carevis: Integrated visualization of computerized protocols and temporal patient data. *Artificial Intelligence in Medicine*, 37(3):203 – 218, 2006.

[3] D. Boaz and Y. Shahar. Idan: A distributed temporal-abstraction mediator for medical databases. In M. Dojat, E. Keravnou, and P. Barahona, editors, *Artificial Intelligence in Medicine*, volume 2780 of *Lecture Notes in Computer Science*, pages 21–30. Springer Berlin Heidelberg, 2003.

[4] P. Bodesinsky, P. Federico, and S. Miksch. Visual analysis of compliance with clinical guidelines. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies*, i-Know '13, pages 12:1–12:8, New York, NY, USA, 2013. ACM.

[5] M. Chen, D. Ebert, H. Hagen, R. Laramee, R. Van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. *Computer Graphics and Applications, IEEE*, 29(1):12–19, Jan 2009.

[6] P. Federico, S. Hoffmann, A. Rind, W. Aigner, and S. Miksch. Qualizon graphs: Space-efficient time-series visualization with qualitative abstractions. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, AVI '14, pages 273–280, New York, NY, USA, 2014. ACM.

[7] T. Gschwandtner, W. Aigner, K. Kaiser, S. Miksch, and A. Seyfang. Carecruiser: Exploring and visualizing plans, events, and effects interactively. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis), 2011*, pages 43–50, March 2011.

[8] S. Miksch, R. Kosara, Y. Shahar, and P. D. Johnson. Asbruview: Visualization of time-oriented, skeletal plans. In *Proceedings of the Fourth International Conference on Artificial Intelligence Planning Systems*, pages 11–18. Association for the Advancement of Artificial Intelligence, 1998.

[9] S. Miksch, Y. Shahar, and P. D. Johnson. Asbru: A Task-Specific, Intention-Based, and Time-Oriented Language for Representing Skeletal Plans. In E. Motta, F. van Harmelen, C. von Pierret-Golbreich, I. Filby, and N. Wijngaards, editors, *7th Workshop on Knowledge Engineering: Methods & Languages (KEML-97)*, 1997.

[10] M. Peleg. Computer-interpretable clinical guidelines: A methodological review. *Journal of Biomedical Informatics*, 46(4):744 – 763, 2013.

[11] M. Peleg, S. Keren, and Y. Denekamp. Mapping computerized clinical guidelines to electronic medical records: Knowledge-data ontological mapper (kdom). *Journal of Biomedical Informatics*, 41(1):180 – 201, 2008.

[12] M. Peleg, Y. Shahar, and S. Quaglini. Making healthcare more accessible, better, faster, and cheaper: the mobiguide project. *European Journal of ePractice*, (20):5 – 20, November 2013.

[13] S. Quaglini. Compliance with clinical practice guidelines. In A. ten Teije, S. Miksch, and P. Lucas, editors, *Computer-based Medical Guidelines and Protocols: A Primer and Current Trends*, volume 139 of *Studies in Health Technology and Informatics*, pages 160–179. IOS Press, 2008.

[14] A. Rind, T. D. Wang, W. Aigner, S. Miksch, K. Wongsuphasawat, C. Plaisant, and B. Shneiderman. Interactive information visualization to explore and query electronic health records. *Foundations and Trends in Human-Computer Interaction*, 5:207–298, 02 2013.

[15] Y. Shahar. A framework for knowledge-based temporal abstraction. *Artificial Intelligence*, 90(12):79 – 133, 1997.

[16] Y. Shahar, O. Young, E. Shalom, A. Mayaffit, R. Moskovitch, A. Hessing, and M. Galperin. Degel: A hybrid, multiple-ontology framework for specification and retrieval of clinical guidelines. In M. Dojat, E. Keravnou, and P. Barahona, editors, *Artificial Intelligence in Medicine*, volume 2780 of *Lecture Notes in Computer Science*, pages 122–131. Springer Berlin Heidelberg, 2003.

[17] E. Shalom and Y. Shahar. A graphical framework for specification of clinical guidelines at multiple representation levels. In *AMIA Annual Symposium Proceedings*, volume 2005, page 679. American Medical Informatics Association, 2005.

[18] J. J. Thomas and K. A. Cook. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society Press, 2005.

[19] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann. A survey on interactive lenses in visualization. In *EuroVis-STARs*, pages 43–62. The Eurographics Association, 2014.