Structuring the Strategies for Integrating Visualization into Automated Modeling Techniques for Visual Analytics

Thorsten May
Fraunhofer Institute for Computer Graphics Research
Darmstadt, Germany
thorsten.may@igd.fraunhofer.de

Juergen Bernard
Interactive Graphics Systems Group
TU Darmstadt, Germany
juergen.bernard@gris.tu-darmstadt.de

Joern Kohlhammer
Interactive Graphics Systems Group
TU Darmstadt, Germany
joern.kohlhammer@gris.tu-darmstadt.de

ABSTRACT

The visual analytics process model [3] became the most widely used model in visual analytics research during the past decade. A plethora of approaches have been presented that integrate visualizations and automatic modeling techniques. Yet today, we consistently face the problem to explain approaches to non-VA-experts in terms of properties that are specific to visual analytics. At the theoretical level of the visual analytics process model, visual analytics properties are reduced to a single connection (i.e. between *visualization* and *model*, see Figure 1). This does not suffice to explain the variety of approaches. Explaining visual analytics at a concrete level by using examples is unsatisfying as well. Firstly, it requires a considerable focus on the details of visualization and modeling techniques, which are already covered by other fields. Secondly, this strategy actually masks the underlying ideas of visual analytics that are *independent from analytical technique, implementation, or application.*

We posit a lack of theory between these two levels of abstraction. We expect such a theory (1) to explain why a technique is a visual analytics approach, (2) to explain ("hands-on") where existing techniques can be turned into visual analytics approaches, and (3) to explain the differences and similarities by parsimonious properties that are not being adopted from other fields of research.

The work closest to such a theory is the survey compiled by Bertini and Lalanne [1]. They distinguish and name different patterns to integrate visualization and automated techniques. We took Bertini and Lalanne's results as a starting point to refine its terminology. Their survey does not focus on potential similarities and underlying building blocks. In addition, we were interested if pattern definitions can be based on common, fundamental concepts. To narrow our scope, we focus on visualizations integrated to modeling approaches. We refined the terminology in three steps. Firstly, we were interested if (some) patterns can be defined in terms of a common terminology. Secondly, we extrapolated this terminology to hypothesize about other potential patterns. Finally, we surveyed if these potential patterns actually have been implemented in recent approaches.

The basis for our terminology is a generic decomposition of an algorithmic modeling approach by Fayyad et al [2]. We observed that many approaches can be described in terms of linked visualizations that show different modeling components (Figure 1, center shows one of these patterns). We distinguish patterns by the combination of components that are visualized or even modified by visualizations.

Our patterns are idealized building blocks to support automated modeling techniques with visualization. These patterns help us structuring the visual analytics solution space. Furthermore, they expose the leverage points to improve or modify existing techniques.

C4PVG '16 Baltimore, MD USA

DOI: 10.475/123_4

1. REFERENCES

[1] E. Bertini and D. Lalanne. Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. In *Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration*, VAKD '09, pages 12–20, New York, NY, USA, 2009. ACM.

[2] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–54, 1996.

[3] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, pages 154–175. Springer-Verlag, Berlin, Heidelberg, 2008.

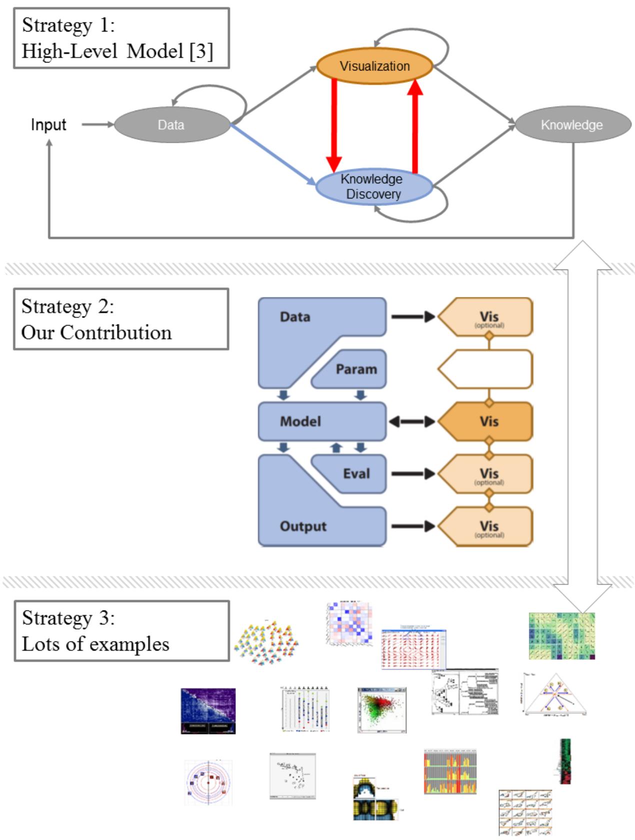

Figure 1: The Visual Analytics Process Model (Strategy 1) represents the abstract idea of integration. Presenting lots of examples (Strategy 3) shows the variety of approaches, but it masks VA concepts that are independent from techniques, implementation, or application. We prepose a terminology on the intermediate level that is able to distinguish patterns of integrating visualization (right) into the modeling process (left).