Computers & Graphics 39 (2014) 37–47

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on Uncertainty and Parameter Space Analysis in Visualization

Supporting the integrated visual analysis of input parameters and simulation trajectories

Martin Luboschik [a,*], Stefan Rybacki [b], Fiete Haack [b], Hans-Jörg Schulz [a]

[a] Computer Graphics Group, University of Rostock, Germany
[b] Modelling and Simulation Group, University of Rostock, Germany

ARTICLE INFO

Article history:
Received 22 June 2013
Received in revised form
27 August 2013
Accepted 28 September 2013
Available online 23 October 2013

Keywords:
Parameter space visualization
Trajectory visualization
Multi-scale visualization
Interactive computational steering

ABSTRACT

The visualization of simulation trajectories is a well-established approach to analyze simulated processes. Likewise, the visualization of the parameter space that configures a simulation is a well-known method to get an overview of possible parameter combinations. This paper follows the premise that both of these approaches are actually two sides of the same coin; since the input parameters influence the simulation outcome, it is desirable to visualize and explore both in a combined manner. The main challenge posed by such an integrated visualization is the combinatorial explosion of possible parameter combinations. It leads to insurmountably high simulation runtimes and screen space requirements for their visualization. The Visual Analytics approach presented in this paper targets this issue by providing a visualization of a coarsely sampled subspace of the parameter space and its corresponding simulation outcome. In this visual representation, the analyst can identify regions for further drill-down and thus finer subsampling. We aid this identification by providing visual cues based on heterogeneity metrics. These indicate in which regions of the parameter space deviating behavior occurs at a more fine-grained scale and thus warrants further investigation and possible re-computation. We demonstrate our approach in the domain of systems biology by a visual analysis of a rule-based model of the canonical Wnt signaling pathway that plays a major role in embryonic development. In this case, the aim of the domain experts was to systematically explore the parameter space to determine those parameter configurations that match experimental data sufficiently well.

1. Introduction

Exploring the parameter space as well as exploring simulation trajectories both struggle with the question, which respective regions should be observed for interesting behavior. And even if these regions were known, it is still open at which level of granularity to look at these regions in order to guarantee that all crucial information is captured. The first question relates to the *extent* of the observation, while the second relates to the *grain* of observation – together they define the observation's *scale* [1, pp. 55–65]. When analyzing the mere outcome of a single simulation, the problem of finding the right scale of observation is already challenging [2]. As one can independently drill down in the two aspects of input parameters and simulation output, their combined exploration increases the search space for a suitable scale dramatically. As a result of this, screen space requirements and computation times increase as well, because it requires to precompute simulations for the multitude of parameter combinations to be able to show them in concert.

In this paper, we contribute a Visual Analytics approach that addresses this challenge and aims to achieve such an integrated visualization of parameter space and resulting simulation trajectories. This approach is based on a coarsely sampled parameter space to address the challenge of high computation times, for which the simulation output is visualized at an even coarser scale to address the challenge of high screen space requirements. To aid in navigating this multi-scale setup, we further provide means to inform the analyst in a coarse-grained overview where it is worthwhile to drill down towards more fine-grained scales. At first, such a drill-down will show the more fine-grained simulation results at lower levels that have already been computed. When the drill-down has exhausted the computed results, it can be used to make an informed decision on whether to run additional simulations for a more fine-grained subsampling. This way, only those regions that actually exhibit behavior of interest are re-simulated at a more fine-grained scale. Through this close tie-in between computation and visual exploration, we effectively avoid an impossible exhaustive simulation, visualization, and exploration of the entire parameter space.

* Corresponding author. Tel.: +49 381 498 7496.
E-mail address: martin.luboschik@uni-rostock.de (M. Luboschik).

0097-8493/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cag.2013.09.004

The following Section 2 will briefly discuss the existing approaches for parameter space visualization, before we will introduce our overall Visual Analytics approach in Section 3. This approach is given as a generic workflow that is applicable as a domain-independent solution strategy and then detailed for our specific context that yields time-series data (1-dimensional trajectories) as outcome of a stochastic simulation. Our realization of this approach is further exemplified by a use case from the domain of systems biology in Section 4. This example investigates an 11-dimensional continuous parameter space for which stochastic simulations with 4096 sampled time steps and 50 replications each are performed. The advantages that our approach has over traditional approaches for inspecting such data, as well as its inherent limitations are discussed in Section 5, before giving some concluding remarks and an outlook on future work in Section 6.

2. Related work

The dependency between input parameters and computation outcome has been subject of visualization research for quite some time. The simplest kind of parameter visualization poses a concrete parameter combination and seeks to visualize the corresponding computation outcome. Examples for such visualizations range from early ones, such as the Influence Explorer [3], to more recent and more sophisticated ones, such as the system described in [4]. While the earlier example features most prominently a number of interactive sliders to set the desired parameter combination, the setting of parameters is already more advanced in the later example, in which variations of control parameters are shown in histograms, scatterplots and alike. Some techniques even integrate the computation outcome into the parameter setting interface as with Image Graphs [5] or spreadsheet interfaces [6].

Computationally and visually more challenging are techniques that have the opposite take on the problem: given a desired computation outcome, they want to determine and visualize combinations of input parameter values that will (approximately) lead to this outcome. To ask the question this way around is a more recent challenge that has led to the development of a number of novel visualization approaches. Among them are result-driven explorations of visual effects [7–9], the accuracy-driven exploration of simulation results [10], and the multi-dimensional optimization of quantification indicators proposed in the Vismon system [11].

All of these approaches have in common, that they aim at individual parameter combinations of simulation results. Thus, visualizations of the entire parameter space that show the simulation outcome for a large number of parameter combinations fall into a third category of parameter visualization. Examples for this category are the parameter space visualization for image analysis [12] and the 2-dimensional embedding of high-dimensional parameter spaces for navigation and interaction [13]. This third category faces the same challenges of high computation time and high screen space requirements as we have outlined them in Section 1. In most cases, visualizations of this category thus apply a sampling of the parameter space as well, in order to cope with these challenges. One solution that enables the user to focus on known regions of interest and prevents him from spending time exploring meaningless parameter combinations is the use of presets, as it has been utilized by [13]. Beyond presets, first approaches like [14] or the Tuner system [15] aim to point the user towards locations in the parameter space that might be of interest but are not yet covered by the current parameter sampling. They use estimation and objective functions to compute these locations, albeit noting that this may not be possible to extend to all application domains – in particular not to those that underlie random processes, like stochastic simulation does.

Drilling down and refining an initial sampling of the parameter space without any estimation – as our approach aims to achieve it – is an application of the domain of *multi-scale visualization*. There are several approaches that aim to provide solutions for investigating data at different scales. These can be divided into approaches that permit their interactive exploration by switching between scales and approaches that mix different scales by partially embedding high resolution data into coarser scales.

The interactive approaches rely on common hierarchical navigation techniques, such as drill-down and roll-up [16], to switch seamlessly between different scales. Sophisticated examples utilizing this navigation scheme are the SignalLens [17], ChronoLenses [18], or the Stack Zooming approach [19]. As such an interactive approach requires some time to fully explore the data, one can employ methods that detect and show regions or points of interest where a drill-down might be worthwhile. For example, the approach of [2] computes the local differences between subsequent scales and depicts them in so-called *heterogeneity bands* alongside the visualization. These alert users where deviating behavior from the currently shown time-plot can be observed by drilling down into scales that are currently below pixel-size and thus not discernable.

The embedding approaches use often very similar measures to the interscale differences described above, but they utilize this information differently by embedding the detected patterns on the finer scale into the coarser scale of the overall behavior of the data [20]. Depending on what patterns these measures detect and subsequently embed, these approaches are called, for example, *outlier-preserving* [21] or *peak-preserving* [22]. Other techniques, such as the Clustering Visualization Spreadsheet [23] try to reintroduce data characteristics that may have been missed due to masking effects back into the visualization.

So far, to the best of our knowledge, both approaches – parameter space visualization and multi-scale visualization – have not been brought together. Yet we see a clear need to do so, as multi-scale visualization can be used to address the challenges that parameter space visualization poses beyond the mere support via presets. They can aid the human analyst in choosing just the right scale to run the simulations at and to look at the resulting outcome, so that computation time and screen space can be allotted to exactly those regions in the parameter space and those parts of the simulation results that show behavior of interest. The next section outlines how we propose to combine these two visualization approaches in order to achieve this.

3. Our principal approach and Visual Analytics setup

Our principal approach consists of three stages which each relate to a different aspect of the Visual Analytics process – the data, the visualization, and the insight gained by the human analyst. This overall approach is described in the following section, before the three stages and their realization in our Visual Analytics setup are introduced in detail in the sections thereafter.

3.1. Our principal approach

Visual Analytics is all about the back and forth between computational analysis, visual exploration, and the human in the loop who is steering this process. In our case, the subject of analysis is a model of a real system and its multi-dimensional parameter space, whereas the computational method of analysis is its simulation. While its purely computational analysis would be too runtime-intensive to be practical, the combination of