

A Survey of Multi-Space Techniques in Spatio-Temporal Simulation Data Visualization

Xueyi Chen^{a,1}, Liming Shen^{a,1}, Ziqi Sha^a, Richen Liu^{a,*}, Siming Chen^{b,c}, Genlin Ji^a,
Chao Tan^a

^a School of Computer Science and Technology, Nanjing Normal University, No.1 Wenyuan Road, Qixia District, Nanjing, Jiangsu, PR China

^b Fraunhofer Institute for Intelligent Analysis and Information Systems, Germany

^c University of Bonn, Germany

ARTICLE INFO

Article history:

Available online 16 August 2019

Keywords:

Simulation data visualization
Spatio-temporal data visualization
Comparative visualization

ABSTRACT

The widespread use of numerical simulations in different scientific domains provides a variety of research opportunities. They often output a great deal of spatio-temporal simulation data, which are traditionally characterized as single-run, multi-run, multi-variate, multi-modal and multi-dimensional. From the perspective of data exploration and analysis, we noticed that many works focusing on spatio-temporal simulation data often share similar exploration techniques, for example, the exploration schemes designed in simulation space, parameter space, feature space and combinations of them. However, it lacks a survey to have a systematic overview of the essential commonalities shared by those works. In this survey, we take a novel multi-space perspective to categorize the state-of-the-art works into three major categories. Specifically, the works are characterized as using similar techniques such as visual designs in simulation space (e.g. visual mapping, boxplot-based visual summarization, etc.), parameter space analysis (e.g. visual steering, parameter space projection, etc.) and data processing in feature space (e.g. feature definition and extraction, sampling, reduction and clustering of simulation data, etc.).

© 2019 Zhejiang University and Zhejiang University Press. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

With the improvement of computing power and the development of supercomputers, simulations in different scientific domains have become increasingly popular. They are widely used in scientific applications, such as climatology, aerodynamics, etc. The simulation data consist of the data in simulation parameter space, simulation process, simulation output, and even the data in in-situ visualization. Most of the scientific simulation data are spatio-temporal. The simulation data are often characterized as multi-run, multi-variate, multi-modal and multi-dimensional.

Spatial-temporal simulation data visualizations are designed to compute and analyze the sensitivity, uncertainty and stability of the simulation models. Specifically, the visual designs are able to reveal the uncertainties in simulation space, parameter setting in parameter space and data distribution in feature space. Studying the whole dynamic process is not only necessary for obtaining a full understanding of the simulation model, but also

expected to reveal more information about the evolution of the information in the model.

1.1. Related surveys

There is only one general review on visualization and visual analytics works for multifaceted scientific data (including multi-run simulation data) by Kehrer and Hauser (Kehrer and Hauser, 2013) in 2013. Considering a broader view of related surveys, there are only parts of the research works mentioned in the survey of research challenges in ensemble data (Obermaier and Joy, 2014; Wang et al., 2015), a survey of data reduction techniques for simulation data (Li et al., 2018), and the state-of-the-art work in meteorology simulation (Rautenhaus et al., 2017). They either described a small number of representative papers in detail, or lacked of summarization on those essential commonalities and a systematic overview. We believe that a complete survey of the state-of-the-art work with a novel and essential perspective is necessary.

We searched for relevant papers to add them into this survey on the IEEE Xplore, EuroGraphics Digital Library and the ACM digital library. The papers are collected from IEEE TVCG, Computer Graphics Forum, IEEE Visualization, EuroVis, IEEE PacificVis, etc. Then we kept the papers using one of multi-space visualization techniques or one of the multi-space analysis techniques.

* Corresponding author.

E-mail address: richen@pku.edu.cn (R. Liu).

¹ The former two authors contribute to this work equally.

Peer review under responsibility of Zhejiang University and Zhejiang University Press.

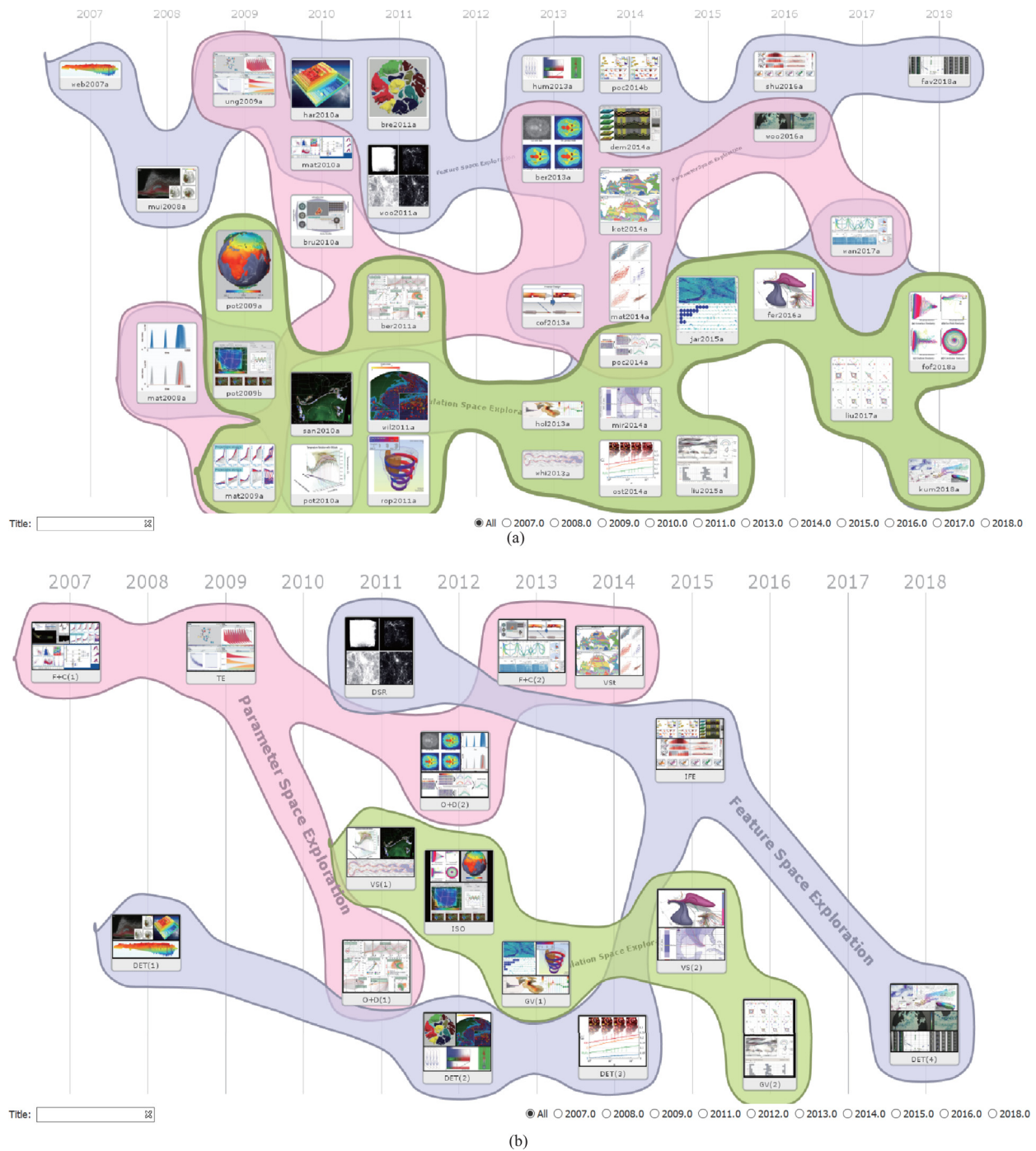


Fig. 2. An overview of rmmallmost of the papers in 10 sub-categories visualized by bubbles (Collins et al., 2009). We improved this interactive real-time tool to make different levels of categories of the literature clear and easy to understand. The tool can be used to explore the literature in a focus+context way. Each node represents a paper (a) or a sub-category (b) using identical techniques. Each node can be enclosed by bubbles to show multiple techniques it used, while the colors of the bubbles represent different techniques. The horizontal axis represents the publication time while the vertical axis represents the citation data collected from ScienceDirect (www.sciencedirect.com) and Google Scholar (scholar.google.com).

statistics plot to show the joint average of humidity and temperature across altitude slices. In each slice, the joint mean, joint standard deviation, skew variance and covariance are visualized to show the summary statistics, as shown in Fig. 3(b). Besides, the joint data histogram are shown as jittered quadrilaterals. In this work, uncertainty is conveyed by not only the traditionally-used variables, i.e., mean, standard deviation, skew variance and covariance, but also the joint data histogram, which is used to show density information. All the variables are visualized by

boxplot and well-designed glyphs in simulation space. Clustering is also used to summarize different data patterns in feature space.

Similar to the work of the summary plot (Potter et al., 2010), a visualization tool named Noodles (Sanyal et al., 2010) was designed to interactively summarize the ensemble simulation output and associated uncertainty of climate simulation data by spaghetti plot and circular glyphs in simulation space. Spaghetti plot is rendered to show the mean and standard deviation and quartile, as shown in Fig. 3(a). However, when many features overlap, spaghetti plots in Noodles may suffer from too much

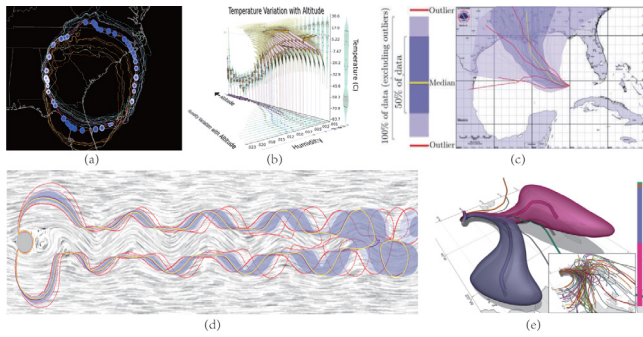


Fig. 3. Boxplot-based visual summarization. (a) A summary view named Noodles (Sanyal et al., 2010) was proposed to visualize uncertainties presented in ensemble simulation data. (b) Two orthogonal 2-D views to form a 3-D view to show 2-D summary plot of uncertainty visualization (Potter et al., 2010). The joint average of humidity and temperature across altitude slices. The joint representations of mean, standard deviation, skew variance and covariance are shown in the glyph design. (c) A curve boxplot visualization of an ensemble of 50 simulated hurricane tracks (Mirzargar et al., 2014). (d) Contour boxplot for an ensemble of the pressure field of a fluid flow simulation with a line integral convolution background image for context (Whitaker et al., 2013). (e) An abstract visualization of the major trends in 2-D and 3-D (Ferstl et al., 2016).

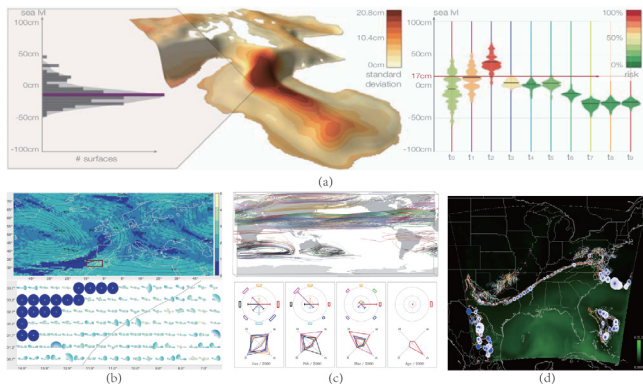


Fig. 4. Glyph-based approaches to conducting comparative visualizations. (a) The first linked view (left) shows a histogram of depth positions of the surfaces at a selected spatial position and time step. The second linked view (right) is a time-series view that depicts a glyph for each time step at the selected position (Höllt et al., 2013). (b) Glyph-based techniques allowing an interactive comparative exploration of 2-D ensembles of vector fields (Jarema et al., 2015). (c) Glyphs in small multiples were designed to compare vortex shape and vortex size across different ensemble vector fields (Liu et al., 2015, 2017). (d) Noodles (Sanyal et al., 2010): visualization of water-vapor mixing ratio to illustrate the progression of uncertainty by the use of spaghetti plots and uncertainty glyphs.

visual clutter, so that they cannot easily convey the main trends, outliers, and statistical properties of the feature distribution. Then, to overcome the limitations, several boxplot-based works are further proposed, namely, contour boxplot (Whitaker et al., 2013) and curve boxplot (Mirzargar et al., 2014). The former one (contour boxplot) presents a generalized method on extending functional data depth to contours and demonstrated methods for displaying the resulting boxplots for 2-D climate simulation data and 2-D computational fluid dynamics, as shown in Fig. 3(d). The latter one (curve boxplot) proposes a generalized method to extend traditional whisker plots or boxplots to curves traced in ensemble vector fields. Especially, the curve boxplot employs a nonparametric method to summarize ensemble pathlines. It is an extension of a method from descriptive statistics (Noodles) and data depth (contour boxplot) to multidimensional curves (curve boxplot), as shown in Fig. 3(c).

Furthermore, a more flexible tool, i.e., streamline variability plots (Ferstl et al., 2016), was proposed for visualizing the statistical properties and uncertainties of multiple aggregation trends of streamlines in simulation space. Technically, they used principal component analysis (PCA) to transform a set of streamlines into a low-dimensional Euclidean space, and further used the principal component representation to depict a new concept, i.e., streamline median. It represents 2-D/3-D regions with high confidence along time. Each streamline median can be extracted from one aggregation trend of streamlines. The relative strength of a trend is depicted by the thickness of its median line and the boxplot. The 2-D and 3-D results are shown in Fig. 3(e). Most of the boxplot series works used clustering algorithm to get clusters or outliers and analyze them in feature space. Besides, Liu et al. (2016) proposed a LCSS-based measurement to compute similarity field and uncertainty field among ensemble pathlines.

2.2. Glyph-based visual mapping

To address the challenges posed by spatio-temporal simulation data visualization, glyph-based approach is a reasonable choice for visual mapping. Glyphs are characterized as symbolic or iconic representations of one or more variables of a data set. Their visual representation can be changed by altering the glyph properties (Ropinski et al., 2011). In this section, we review glyph-based techniques applied in simulation data visualization. Well-designed glyphs could be used to visualize the uncertainty, sensitivity, or stainability presented in the whole simulation processes.

Sanyal et al. (2010) designed multiple views of ribbon and glyph-based uncertainty visualization, isocontour colormaps and spaghetti plots. The circular glyphs scaled in size are one of the most popular techniques to measure uncertainty in scalar field ensembles. The maximum size of the glyphs is limited by the spacing such that the overlap between glyphs is minimized. Standard-deviation, the width of the 95% confidence interval and interquartile range were encoded into the size of circular glyphs, as shown in Fig. 4(d). The glyphs of concentric circles represented the confidence interval and interquartile range while the user could select bootstrap mean or ensemble mean across different simulation runs.

Except for the scalar field ensemble visualization, glyph is often employed to vector field ensemble visualization. For example, a glyph-based technique (Jarema et al., 2015) was proposed to perform interactive comparative visualization on 2-D vector field ensembles. It encodes the variation modes into shapes and directions of glyphs, as shown in Fig. 4(b). Nevertheless, this comparison was limited to a 2-D scenario. Liu et al. (2015, 2017) designed some glyphs to make a metaphor for user-defined derived features (i.e., vortex) from vector field ensembles. Vortex is a significant feature in vector field such as climate simulation data and ocean simulation data. Then they compare derived features by glyphs placed in small multiples in 3-D simulation space, as shown in Fig. 4(c).

Besides, glyph is often used to make a metaphor in simulation space, for example, glyph is placed at the selected location in each time step in times-series data to encode the risk of sea depth and further provide interactive visual analysis in ocean surface simulation data (Höllt et al., 2013). For a more detailed description of the entire distribution of the surfaces output in multiple simulation runs, they provided two linked views. The left linked view shows the depth position histograms of the simulated sea surface, while the right linked view is a glyph representation view. The horizontal axis shows the selected critical sea level, where the color of each glyph encodes a risk whose distribution is above the critical level (Fig. 4(a)). Except for the metaphor made in sea level risk in ocean simulation, glyph can reveal the number

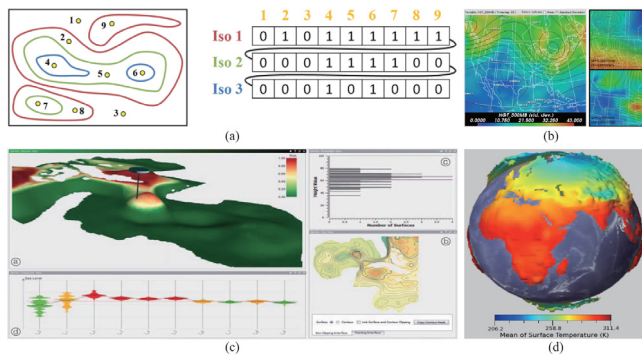


Fig. 5. (a) Extending from one to three isocontours (Fofonov and Linsen, 2018). (b) Color plus overlaid contours (left) and close ups which both show the same regions of data (right) (Potter et al., 2009a). (c) Users are allowed to specify a critical height value, whose iso-contour derived from the currently selected surface is then highlighted (Höllt et al., 2013). (d) Height fields of the mean of the surface temperature (Potter et al., 2009b).

of variation modes in ensemble simulation and the spread across the whole simulation runs. Interactions are designed to explore spatio-temporal simulation data in simulation space and feature space.

2.3. Silhouette-based illustrative rendering

Silhouette-based visualization techniques can illustrate the field data (i.e., the data in the scalar field, vector field, tensor field, etc.) in simulation space, which help to reveal the patterns and distribution of the data. Some frequently-used approaches contain isocontour in the 2-D field and isosurface in the 3-D field.

Isocontour and isosurface should be extracted before the visualization and exploration. There had been a series of work applying statistical methodology directly onto isocontour detection in spatio-temporal simulation data. Höllt et al. (2013) developed a visualization system according to different requirements of different stages. It allows users to specify a critical height value, whose iso-contour extracted from the selected surface is then highlighted, as shown in Fig. 5(c). Besides, in order to explore the parameter space, such as viewing the influence of a certain parameter or removing outliers, the isocontour extraction and statistical analysis can be carried out either for the complete ensemble, or for any user-defined subset of the ensemble.

Furthermore, Potter et al. (2009b) presented an approach to explore ensemble simulation data by using a set of statistical descriptors to summarize the data, visualizing these descriptors using various of visualization techniques such as multiple 2-D plots, slicing and 3-D iso-contouring, as shown in Fig. 5(d). In the same year, Potter et al. (2009a) proposed a framework named Ensemble-Vis which consists of a set of overviews and detailed statistical views and provided some interactions. They used a variety of techniques about isocontour such as the spaghetti plot which allows to choose values for different simulation runs of the ensemble. In contrast to the approaches presenting diverse information in a single display, combining multiple linked views yields a better presentation of the data and facilitates a larger level of visual data analysis, as shown in Fig. 5(b).

In order to increase the accuracy of the measure, users expert to choose more than one iso-contour simultaneously. Fofonov and Linsen (2018) generalized the isosurface similarity to a field similarity and further to a multi-field similarity. They are used to catch local field differences between different data frames instead of just providing aggregated statistical information, as shown in Fig. 5(a). Most of the silhouette-based methods often employ data reduction methods (in feature space) to summarize the isosurface similarity or difference across simulation runs for comparison.

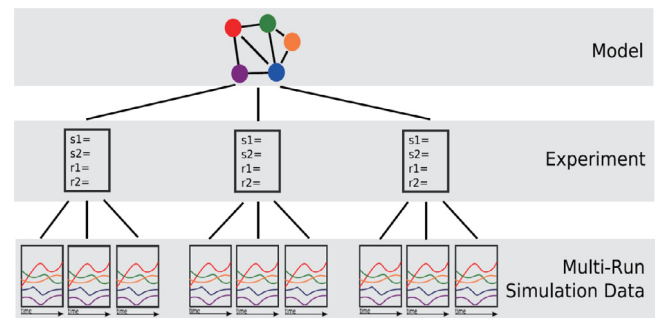


Fig. 6. Trial-and-error experiment parameter assignments in parameter space exploration. Each experiment result corresponds to a different set of parameters (Unger and Schumann, 2009).

3. Parameter space exploration

For the purpose of getting a convincing conclusion, domain scientists often use the same simulation model with various values of control parameters, generating large data sets, i.e. state parameters, which capture different aspects of the behavior of the target phenomenon (Matkovic et al., 2010).

Generally, users should make use of their intrinsic knowledge about the process and the visual information on the screen (Unger and Schumann, 2009). Considering that the inputs could be multivariate, high-dimensional parameter space exploration could be tedious and complicated even for the experts. In this section, we discuss the works based on the exploration of parameter space, including parametric visual exploration and visual steering. The former one can be further categorized into three subcategories of techniques, including trial-and-error, focus+context, and overview-to-detail. Similarly, many summarized papers might belong to multiple different categories simultaneously.

3.1. Parameter visual exploration

Experts often need to explore the parameter space so that they could identify a more appropriate region for the inputs for the subsequent simulation runs, gaining more accurate simulation outputs. A visual representation of the model captures the dependencies, allowing engineers to speculate the relationships between the behaviors of the models and parameters with their domain expertise in the scenario of multiple simulation runs. We could derive three main studies focusing on this process of the exploration: trial-and-error, focus+context and overview-to-detail.

3.1.1. Trial-and-error exploration

The most straight-forward way is to simulate with the same model and different sets of parameters for many times and show the results so that experts could narrow the input space down to get a better response generated from the model. Unger and Schumann (2009) aimed at providing visual support for the understanding of underlying simulation processes and the abstract data sets, which needs the visual combination of both of them. In this way, the efforts for gathering all the necessary information for decision-making are reduced. They analyze and compare multi-run and multi-variate time series on three different process levels: the model, experiment and the multi-run. Multiple experiment views are coordinated at the model level, providing a comparison of the experiments of a model. Different experiment results correspond to different sets of parameters, as shown in Fig. 6. Users can assign the parameters in a trial-and-error way. Eventually, the experiment view is used as an overview at the level of multi-run simulation data.

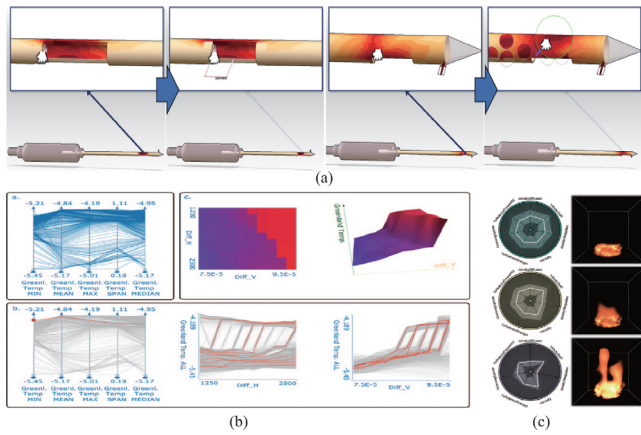


Fig. 7. Focus+context exploration techniques used in spatio-temporal simulation data visualization. (a) Drag operations change the value of simulation inputs in forward design (left), or the simulation outputs in inverse design (right) (Coffey et al., 2013). (b) Different ways of depicting a family of surfaces: a single polyline in the parallel coordinates showing one selected surface, and the 3-D surface view and the 2-D height map view of the selected surface (Matkovic et al., 2009). (c) Parameter view for three clusters of flame simulations, and renderings of representative cluster members (Bruckner and Möller, 2010).

3.1.2. Focus+context exploration

It is worth mentioning that the papers summarized in this sub-section are just to use F+C techniques in parameter space exploration. Focus+context (F+C) visualization could handle large data sets with the nature of high dimensionality, guiding the users and supporting interactive analysis. Matkovic et al. (2009) proposed a parallel coordinates plot (PCP)-based parameter space exploration technique that enables a thorough investigation of simulation outputs, i.e., the families of data surfaces. Users are allowed to brush the parameter space in PCP by using a F+C scheme, as shown in Fig. 7(b).

As the model complexity grows, it becomes incredibly hard to figure out the dependencies between the simulation model behavior and the data sets. Matkovic et al. (2010) further proposed a simulation model view to bridge this gap. The view provides a 2-D graph where each node represents a generation block of the simulation model. In each node, the values of both the control parameters that are used to tune the simulation and state parameters which capture the behaviors of the model are displayed directly. For multiple runs, they define a family of curves and one curve for each run. The curve view combined with multiple linked views and composite brushing is designed to display curves in focus and those curves form the context (F+C).

Besides, many physically-based simulation software tools, which are designed for generating realistic animations, provide no visual guidance in the parameter selection process, making users have to resort to a time-consuming and cumbersome trial-and-error strategy. To solve this issue, Bruckner and Möller (2010) presented a result-driven visual approach for the visual exploration of parameter spaces. The input parameters of simulation can be achieved by an interactive result-driven F+C method, as shown in Fig. 7(c). The system samples the parameter space and employs a novel approach for clustering the resulting volumetric time sequences in order to figure out characteristic variations related to their temporal evolution.

The problem of visualizing sets of simulation results is an active area of research with recent applications to engineering design. Different from small multiples and multiple linked

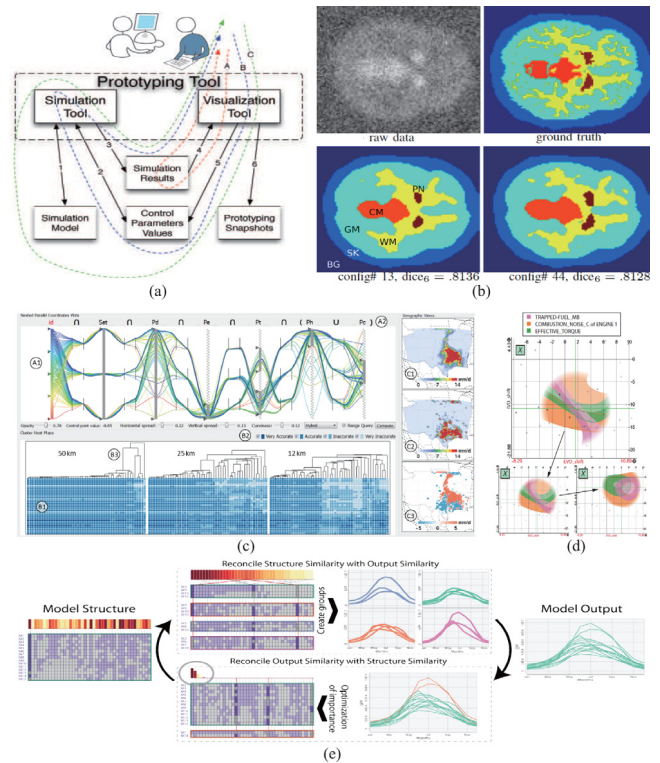


Fig. 8. Approaches for overview-to-detail exploration. (a) Three loops A, B, C are used to distinguish three levels of the interactive steering process (Matkovic et al., 2008). (b) Images in the first line show raw image data from a brain scan, and the ground truth segmentation respectively. The second line shows two output segmentations from the model under two different parameter configurations (Bergner et al., 2013). (c) Visual analytics system designed for multi-resolution climate ensemble parameter analysis with nested parallel coordinates plots (Wang et al., 2017). (d) Mapping the neighborhood of three targets into X, showing the sensitivities to combined variations of X_i and X_j (Berger et al., 2011). (e) Iterative visual reconciliation of groupings based on climate model structure and the output (Poco et al., 2014a).

views (Doleisch et al., 2003), Coffey et al. (2013) argued that the directness of the interface makes it possible to do effective ensemble visualization in-place, which means using essentially a single-view interaction and visualization space rather than visual layouts based on small multiples and multiple linked views. They designed an interface for exploring large exploration spaces as encountered in the tasks that require tuning parameters of computationally-intensive simulations, as shown in Fig. 7(a).

To help domain scientists understand the connection between multi-resolution convective parameters and the large spatial-temporal ensembles, Wang et al. (2017) proposed an integrated visualization system, presenting intricate climate ensembles with a comprehensive overview and on-demand geographic details. Different from the idea of overview-to-detail, domain scientists are more interested in visually quantifying the difference resulted from parameter changing. To tackle the challenge of visualizing multi-resolution high-dimensional parameter space, they propose a new type of parallel coordinates plots named nested parallel coordinates plot, which enables visualization of intra-resolution and inter-resolution parameter correlations of different parameters, as shown in Fig. 8(c). Heat maps and dendrograms are linked together to help domain scientists gain an overall understanding of large multi-resolution ensembles. Additionally, multiple side-by-side geographic views are provided to show spatial and temporal details on demand.

3.1.3. Overview-to-detail exploration

A common goal in visualization is the design of techniques that provide both overview visualizations and support for detailed feature exploration. Researchers often prefer to make designs that present an outline of the data first, just follows the principle: “Overview first, zoom and filter, then details on demand”. The overview usually assists the experts into grasping the overall situation so that they could make a quick decision about narrowing the potential regions down. Generally, it is impossible to run all possible simulation runs at the beginning and analyze the results due to the complexity of the model, which could cause unnecessary wasting of time and computation resources. In this case, [Matkovic et al. \(2008\)](#) proposed a tightly coupled steering loop with interactive visual analysis. They start from a non-optimized initial prototype and the corresponding simulation model and go through an iterative process in detail, as shown in [Fig. 8\(a\)](#). The prototyping continues through the refinement of the simulation model, of the simulation parameters and through trial-and-error attempts to an optimized solution, going back and forth between three different levels of abstraction (simulation data, control parameters values, simulation model).

Besides, a visualization system named ParaGlide ([Bergner et al., 2013](#)) also follows the design principle: “Overview first, zoom and filter, then details on demand”. Specifically, the ParaGlide is designed for interactive exploration of parameter spaces (overview) of multi-dimensional simulation models, which endeavors to facilitate the process of refining models by guiding data generation using a region-based interface for parameter sampling (details), dividing the model’s input parameter space into partitions that propose distinct outputs. It also allows for clustering the output space into groups of similar model results manually, as shown in [Fig. 8\(b\)](#).

For some intricate cases, such the exploration of high-dimension parameter space, experts usually need to make use of the parameter space projection so that they could do dimension reduction with a large number of complex computations to the parameter space before the process of simulation. [Poco et al. \(2014a\)](#) proposed an iterative, human-in-the-loop refinement strategy for reconciling alternative similarity spaces, which leverages the high bandwidth of human perception system and exploits the pattern detection and optimization capabilities of computing models. Since the scientists need to understand the dependency relationships between model structure similarity and model output similarity, this work stems help them to understand the dependency between alternative similarity spaces for climate models, facilitate iterative refinement of groups with the help of a feedback loop. It also allows flexible multi-way interaction and exploration of the parameter space for reconciling the importance of the model parameters with the model groupings ([Fig. 8\(e\)](#)).

Similarly, [Berger et al. \(2011\)](#) presented an interactive approach to enable a continuous analysis of a sampled parameter space with respect to multiple target values. They employ methods from statistical learning to predict results in real-time at any user-defined point and its neighborhood, and describe techniques to guide the user to potentially interesting parameter regions in the context of multiple target dimensions and other application-specific criteria, and visualize the inherent uncertainty of predictions in 2-D scatterplots and parallel coordinates, as shown in [Fig. 8\(d\)](#).

In brief, the techniques about parameter visual exploration could be summarized as trial-and-error ([Matkovic et al., 2010](#); [Unger and Schumann, 2009](#)), which could be applied when the complexity of the parameter space is relatively low, overview-to-detail, which has been accepted widely according to the information seeking mantra, and F+C, which could be more appropriate

when detail is top priority. No matter F+C or overview-to-detail, most of these literatures had designed interactive explorations schemes in feature space to verify and analyze the simulation outputs.

3.2. Visual steering

Considering that it could be anything but intuitive for experts to make decisions by simply observing the visualization results generated by parameter visual exploration systems, an interactive system is usually required to support parameter visual refinement and assist in the analysis. Visual steering is an intuitive and interactive scheme to help domain experts explore the parameter space in spatio-temporal simulation data.

Visual steering approaches can be divided into two subcategories. First, experts could manipulate the outputs of the models directly. By dragging the items in the simulation space or changing their geometry properties, it is convenient to refine the inputs reversely. Second, experts can also establish a surrogate model, such as regression model, to do some adjustments indirectly.

The system designed by [Coffey et al. \(2013\)](#), concentrating mainly on F+C visual steering, integrates forward design via direct manipulation of simulation inputs in the same visual space with inverse design via “tugging” and reshaping simulation outputs, as shown in [Fig. 7\(a\)](#). Also, [Bruckner and Möller \(2010\)](#) presented a work to generate the results in an interactive visual steering environment, which combines 3-D animated views with an abstracted representation of the identified spatio-temporal clusters, as shown in [Fig. 7\(c\)](#). Besides, an overview-to-detail design ([Matkovic et al., 2008](#)) was integrated into a visual steering method in a prototyping environment for automotive industry system design. The three levels of iteration (simulation data, control parameters values, simulation model) provide different levels of interactivity, as shown in [Fig. 8\(a\)](#).

To broaden the scope of the analysis, reduce subjectivity, and facilitate comparison of model output with reference data, [Kothur et al. \(2014\)](#) proposed a novel visual steering approach. It allows modelers to create multiple spatial clusterings of the temporal profiles in model output and reference data, consolidating the various clusterings for each data set with an ensemble approach, and interactively explore and compare the consolidation results, as shown in [Fig. 9\(a\)](#).

Sometimes, experts are willing to establish a surrogate model between the inputs and the results, by analyzing which they could get the information like relativity and sensitivity of the parameters and the outputs. The regression model would be a typical choice for this situation. [Matkovic et al. \(2014\)](#) introduced a novel approach to hybrid visual steering of simulation ensembles. They combine interactive exploration and analysis with automatic optimization based on regression models. An adaptive exploration of simulation space and parameter space is introduced by using a regression model and the use of optimization to find an optimum within a subset of the space, as shown in [Fig. 9\(b\)](#). This novel scheme covers the spectrum between a fully automatic simulation and the manual adjustment of simulation parameters.

In conclusion, the essences contained within visual steering is that it is a possible channel for experts to get a better understanding for the relationship between the input parameters and outputs generated by the model, paving the way for further adjustment of the parameters as well as the models. The techniques about visual steering could be summarized as direct visual steering, which could be intuitive, and indirect visual steering with a surrogate model. Most of these methods use interactive techniques to explore feature space and then verify the simulation output results.

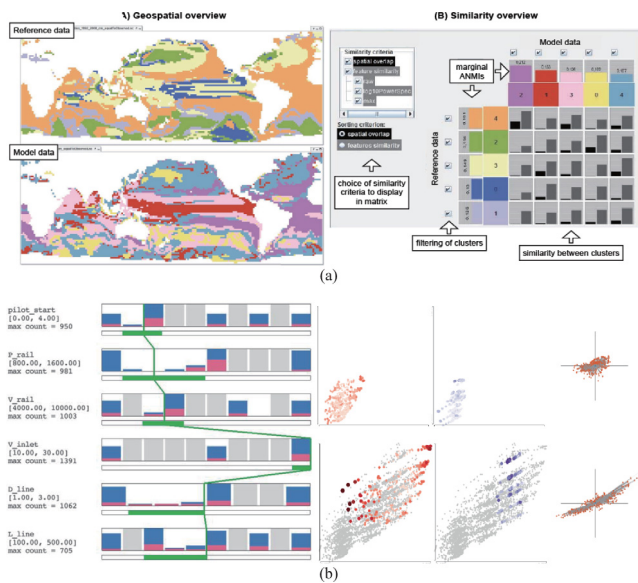


Fig. 9. Visual steering techniques used in spatio-temporal simulation data visualization. (a) The consolidation overview component. Users can choose the parameters, i.e., the similarity criteria to compare the model data and reference data (Kothur et al., 2014). (b) The left view shows six parameters with constraints and optimum values. The right view is designed to compare the simulation results (in orange) and the regression model results (in blue). (Matkovic et al., 2014).

4. Feature space exploration

To improve the accuracy and efficiency of the models in numerous practical fields, such as ocean current, climate forecast, combustion properties, etc., scientists are in the face of analyzing and processing the extremely large data sets generated by repetitive simulation experiment and various parameters. Therefore, in order to reduce data sets and efficiently interact with users, many new techniques are well designed to extract specific feature. The major challenges and relevant techniques will be discussed in three aspects, feature specification, data processing and reduction and interactive feature exploration.

4.1. Feature definition, extraction and tracking in simulation data

Before exploring feature space in simulation processes, it is necessary to define different features based on the multi-run, multi-variate and multi-time simulation data. Some derived feature needs to be further extracted by well-designed algorithms or approaches. Visualization tends to focus on essential parts of the data sets instead of showing all the information and details at the same time. Considering that visualizing high-dimensional and large data sets is no doubt a challenge for users, feature-based exploration is useful to reduce the data sets by well-designed algorithms and help users extract and specify the most interesting features.

In this section, we will discuss some representative approaches of feature extraction in simulation processes, such as the use of feature definition language (FDL), time-series feature specification, feature surface extraction, topological feature analysis and uncertainty feature extraction, as shown in Figs. 10 and 11.

In order to better define the feature, a feature definition language (FDL) (Doleisch et al., 2003) can be designed. It allows the definition of one or several features, which can be complex and hierarchically described by brushing multiple dimensions, as shown in Fig. 10(a). FDL was further used by Muigg et al. (2008) to reduce visual clutter and occlusion. This work defines feature

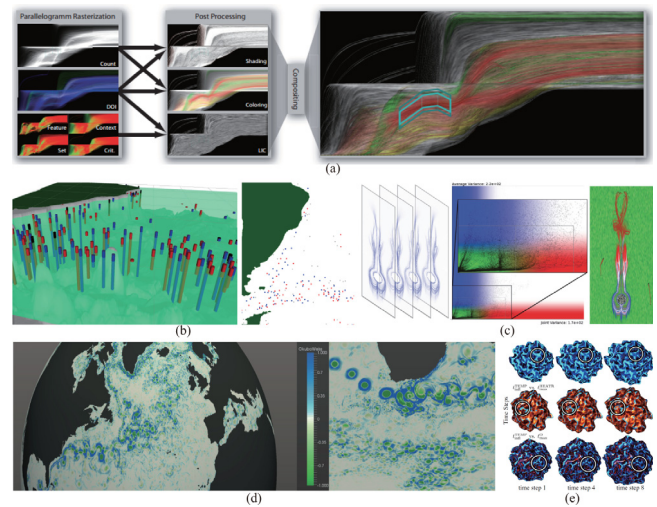


Fig. 10. Typical approaches of feature space exploration. (a) A feature definition language (FDL) is designed to describe features by brushing multiple dimensions (Muigg et al., 2008). (b) The domain specific feature (eddy) is extracted from ocean simulation data (Williams et al., 2011). (c) Ensemble uncertainty is measured through individual variance and joint variance (Hummel et al., 2013). (d) The domain specific feature (eddy) is extracted from ocean simulation data (Woodring et al., 2016). (e) Feature surfaces (Oster et al., 2014) can be extracted to analyze premixed combustion simulation data without reconstructing the data on the original grid.

using FDL when interactively visualizing scientific data by four level method of focus + context, including three different kinds of focus integrated by multiple linked views and the context to get feature specification.

Time-series data are another significant feature in simulation data visualization. Unger and Schumann (2009) used three process levels, i.e., model, experiment and the multi-run simulation data, to design the visualization based on both underlying process and the needs of the users. The visualization of the model level coordinates multiple instances for the comparison of experiments view, while the multi-run simulation data view provides an overview of the data, whose details can be analyzed in time series.

Feature extraction approaches are often employed in ocean simulation data, whose domain specific features are eddies. Williams et al. (2011) presented an analysis of flow data to show the three-dimensional structure and distribution of ocean eddies, as shown in Fig. 10(b). The characteristics of each eddy are recorded to form an eddy census that can be used to investigate correlations among variables of eddies, such as thickness, depth, and location. However, the method is limited by availability and fidelities when data sets are too large to handle. Thus, Woodring et al. (2016) developed an in-situ eddy census workflow. Its green regions indicate strong rotations and blue regions indicate strong shear, while the boundary between regions indicates an eddy, as shown in Fig. 10(d).

A feature surface (Oster et al., 2014) was extracted to analyze numerical simulation data of premixed combustion without reconstructing the data on the original grid. It uses mesh extraction to especially handle the enormous size of the DNS data of premixed combustion, as shown in Fig. 10(e). Their works describe the original data by the flame surface mesh, a sparse data representation, and reconstruct them to the dense scalar fields. And through the different distance of each mesh vertex, one can analyze and compare the behavior of the scalar fields.

Uncertainty, generated from the process of ensemble simulation, can be visualized to analyze the ensemble flow behaviors

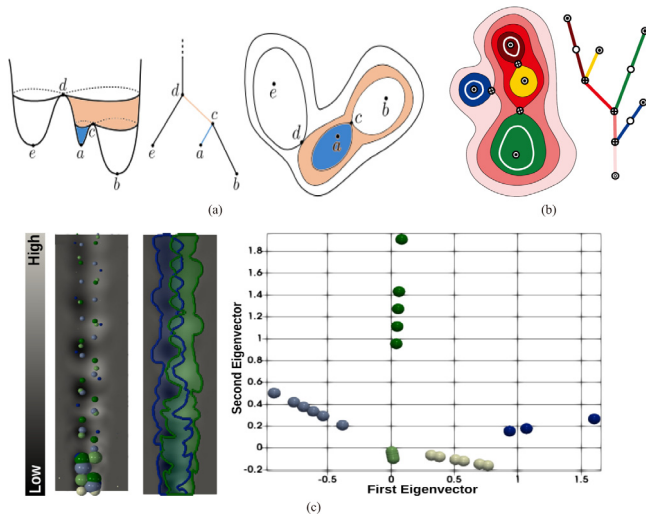


Fig. 11. Topological feature extraction and analysis. (a) Scalar functions are defined on simply connected domains of arbitrary dimension based on the topological landscapes metaphor (Harvey and Wang, 2010). (b) A hierarchical merge tree is constructed by recording the merging of contours for each time step (Bremer et al., 2011). (c) The persistence atlas (Favelier et al., 2019) analyzed the structure of the ensemble in terms of critical point layouts and provided low dimensional embedding of the members, and automatically identified distinct trends in critical point layouts.

in feature space. Hummel et al. (2013) developed a Lagrangian-based uncertainty measurement on vector field ensembles. It connects and transports characteristics of individual Lagrangian neighborhoods to joint flow behavior presented in the ensemble. Thus, both individual variance and joint variance are visualized in feature space and further employed to measure the uncertainty feature in CFD simulation data, as shown in Fig. 10(c). Kumpf et al. (2018) proposed a novel workflow to improve the accuracy of meteorology forecast. It computes the sensitivity of a scalar forecast quantity in simulation data, locating a field where the forecast errors originated, and thus, improves the forecast model.

Essentially, feature tracking algorithms can be divided into two main categories: tracking by geometry and tracking by topology (Favelier et al., 2019). The former calculates distance among geometric attributes, while the latter topologically computes tracking simulation data using contour trees, reeb graphs and critical points, etc. Harvey and Wang (2010) visualized scalar functions defined on an ensemble of terrain models and completely collected the identical contour trees and topological persistence to the input scalar field, as shown in Fig. 11(a). They used a new metaphor called “topological landscapes” (Weber et al., 2007), whose basic idea is to construct a terrain with the same topology as a given data sets and to display the terrain as an easily understood representation of the actual input data.

Hierarchical merge tree and Reeb graph are two commonly-used approaches in topological analysis, while merge tree is further used to reduce and segment data in simulation data analysis when the feature is highly compacted and flexibly represented. A hierarchical merge tree is constructed by recording the merging of contours for each time step (Bremer et al., 2011), as shown in Fig. 11(b). It shows that topological analysis is an appropriate method used for exploring the parameter space to segment, select and track features.

Critical points, robustly extracted by topological methods, have been used in many important applications. However, there are some thresholds issues in critical point extraction. The visualization is extremely costly when many single sets of features are extracted for each set of thresholds, and any change in a threshold

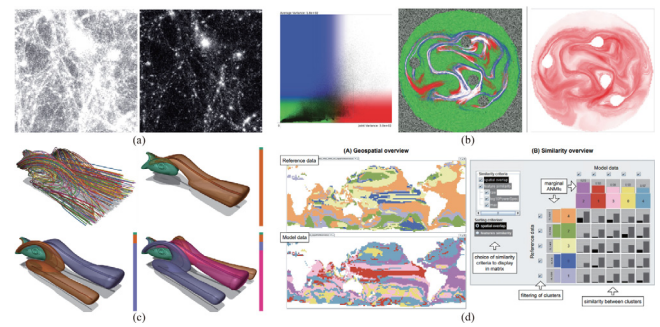


Fig. 12. Feature space analysis of data reduction for simulation data. Data clustering and classification, PCA, and some others. (a) In-situ sampling (Woodring et al., 2011) in large-scale particle simulation: the images show the comparison between full resolution data and the sampled data. (b) PCA is used to compute uncertainty (the linearized deformation or shape change) in ensemble vector fields which were generated by a stirring apparatus simulation (Hummel et al., 2013). (c) PCA is used to convert streamlines into a structure preserving Euclidean space (feature space or PCA space) (Ferstl et al., 2016). (d) Clustering is used to combine all clusters (each one has its unique color) of one data set into a single consolidated clustering (Kothur et al., 2014).

requires the retreatment of data. A commonly-used topological technique named Morse theory addresses it by expressing a similar set of features. The persistence atlas (Favelier et al., 2019) analyzes the structure of the ensemble in terms of critical point layouts, and provides low dimensional embedding of the members, and automatically identifies distinct trends in critical point layouts, as shown in Fig. 11(c). However, there still be a gap that insufficient of the trend variability of critical points. They fill it by providing a 2-D layout of characterized ensembles as critical points to identify and visualize the global trends and outliers by measuring the distance among dissimilarities. Most of the works categorized into the subcategory “feature definition, extraction, tracking” exploit silhouette-based illustrative rendering in simulation space to illustrate the feature or patterns. Actually, isocontour and isosurface can be considered as an important feature in both simulation space and feature space.

4.2. Data sampling and reduction in feature space

Data procession and reduction are essential when comparing the differences and similarities of the aggregation of data sets. The current challenges of reduction are avoiding limited scope, and of losing major factors. It is worth mentioning that the papers summarized in this sub-section are just to use data sampling or data reduction techniques in feature space exploration.

It is wildly spread that using data reduction techniques such as sampling, clustering, classification and dimension reduction can identify the time-varying behavior characteristics, major trends and outliers in data. For example, an in-situ sampling approach (Woodring et al., 2011) was designed to overcome the I/O bandwidth limitation at simulation run-time. In-situ method can reduce the storage bandwidth and stores raw data in post-analysis. The images in Fig. 12(a) show the comparison between full resolution data and the sampled data. A clustering ensemble combines all clusterings of one data set into a single consolidated clustering. A visual interface, as shown in Fig. 12(d), facilitates the comparison of model data and reference data (Kothur et al., 2014).

Principal Components Analysis (PCA) uses the idea of dimensionality reduction to transform multiple indicators into a few comprehensive indicators. Hummel et al. (2013) used PCA to compute dense path line integration, which ensembles properties of the advected neighborhood together, by representing and estimating statistical variance of advected point clouds, as shown

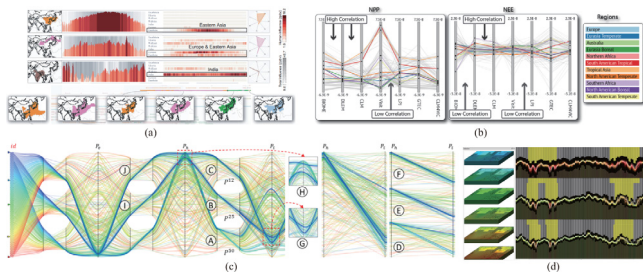


Fig. 13. The techniques combine InfoVis and SciVis schemes to analyze simulation data. (a) Ensemble graph (Shu et al., 2016) was designed to help understand spatiotemporal similarities across ensemble runs. (b) The parallel coordinates plot commonly-used in InfoVis is used to compare the correlations and uncertainties among multiple climate simulation data (Poco et al., 2014b). (c) A nested parallel coordinates plot (InfoVis technique) to visualize the multi-resolution simulation data according to user control in one view (Wang et al., 2017). (d) Multi-charts (Demir et al., 2014): combine line charts and bar charts to show statistical information on ensemble members.

in Fig. 12(b). Furthermore, PCA was used to convert streamlines into a structure-preserving Euclidean space, streamlines can be performed in feature space to detect trends and outliers of clusters (Ferstl et al., 2016), as shown in Fig. 12(c).

4.3. Interactive feature exploration

To extract feature flexibly, there are requires of efficient means of interaction. Users choose the interesting part of data, then specify features extracted correspondingly. Especially driven by the experts working with the visualization system, considerations must be taken when designing one.

Interactive feature exploration is widely used in various practical fields. Linking several views combining information visualization (InfoVis) with scientific visualization (SciVis) techniques can help analyze the simulation data interactively. Shu et al. (2016) used ensemble graph (InfoVis technique) which combined with multiple-linked views showing details to help understand spatiotemporal similarities across runs in time-varying ensemble simulation data (Fig. 13(a)). Multi-charts (Demir et al., 2014) overcomes the occlusion effects in 3-D scalar ensemble fields by using line charts to globally display the correspondences of spacial points, and using the bar charts to display the statistical information of ensemble members, as shown in Fig. 13(d). The parallel coordinates plot is used to compare the correlations and uncertainties among multiple simulation data sets (Poco et al., 2014b), as shown in Fig. 13(b). For each output variable, parallel coordinates enable scientists to analyze the multi-model similarity based on the region-wise distribution of the variable. Wang et al. (2017) used a novel nested parallel coordinates plot (InfoVis technique) to visualize the multi-resolution simulation data, as shown in Fig. 13(c). It visualizes the connection between the large spatial-temporal ensembles. Most of literatures in this category often share the schemes in data sampling and data reduction to reduce visual clutter or decrease the computational complexities and analytical complexities.

5. Discussion and conclusions

Multi-space techniques are summarized in this survey since they are either used individually or synthetically in simulation data visualization, including techniques across multiple space analysis, i.e., simulation space, parameter space and feature space. Spatio-temporal visualization on simulation data analysis can help domain experts or users to get a better insight into the data characteristics, i.e., the data distribution, data pattern, etc. or the

simulation modes. They are used to compute and analyze the model sensitivity, the model uncertainty and the model stability. Therefore, we summarize a systematic overview of the essential commonalities shared by those works.

We notice that most of the simulation space techniques fall into visual design schemes, e.g., visual trend summarization, glyph-based visualization and illustrative rendering by family of surfaces. The parameter space exploration methods can be categorized into trial-and-error, focus-and-context (F+C), overview-to-detail explorations, and visual steering schemes, while the feature space method includes the schemes on feature definition, feature extraction, feature tracking, data sampling, and data reduction.

To our best knowledge, there is no survey papers summarizing the spatio-temporal simulation visualization literatures from the viewpoint of multi-space techniques. Overall, we hope that this survey inspires novel ideas on simulation data visualization by using multi-space techniques.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (NSFC) Grant Nos. 61702271 and 61702270.

References

- Berger, W., Piringer, H., Filzmoser, P., Gröller, E., 2011. Uncertainty-aware exploration of continuous parameter spaces using multivariate prediction. *Comput. Graph. Forum* 30 (3), 911–920.
- Bergner, S., Sedlmair, M., Möller, T., Nabi, S., Abdoljousefi, Saad, A., 2013. Paragliders: Interactive parameter space partitioning for computer simulations. *IEEE Trans. Vis. Comput. Graphics* 19 (9), 1499–1512.
- Bremer, P.-T., Weber, G., Tierny, J., Pascucci, V., Day, M., Bell, J., 2011. Interactive exploration and analysis of large-scale simulations using topology-based data segmentation. *IEEE Trans. Vis. Comput. Graphics* 17 (9), 1307–1324.
- Brockner, S., Möller, T., 2010. Result-driven exploration of simulation parameter spaces for visual effects design. *IEEE Trans. Vis. Comput. Graphics* 16 (6), 1468–1476.
- Coffey, D., Lin, C.-L., Erdman, A.G., Keefe, D.F., 2013. Design by dragging: An interface for creative forward and inverse design with simulation ensembles. *IEEE Trans. Vis. Comput. Graphics* 19 (12), 2783–2791.
- Collins, C., Penn, G., Carpendale, S., 2009. Bubble sets: Revealing set relations with isocontours over existing visualizations. *IEEE Trans. Vis. Comput. Graphics* 15 (6), 1009–1016.
- Demir, I., Dick, C., Westermann, R., 2014. Multi-charts for comparative 3D ensemble visualization. *IEEE Trans. Vis. Comput. Graphics* 20 (12), 2694–2703.
- Doleisch, H., Gasser, M., Hauser, H., 2003. Interactive feature specification for focus+context visualization of complex simulation data. In: *Joint EUROGRAPHICS - IEEE TCVG Symposium on Visualization*.
- Favelier, G., Faraj, N., Summa, B., Tierny, J., 2019. Persistence atlas for critical point variability in ensembles. *IEEE Trans. Vis. Comput. Graphics* 25 (1), 1152–1162.
- Ferstl, F., Bürger, K., Westermann, R., 2016. Streamline variability plots for characterizing the uncertainty in vector field ensembles. *IEEE Trans. Vis. Comput. Graphics* 22 (1), 767–776.
- Fofonov, A., Linsen, L., 2018. Projected field similarity for comparative visualization of multirun multifield timevarying spatial data. *Comput. Graph. Forum* 1–14.
- Harvey, W., Wang, Y., 2010. Topological landscape ensembles for visualization of scalar-valued functions. *Comput. Graph. Forum* 29 (3), 993–1002.
- Höllt, T., Magdy, A., Chen, G., Gopalakrishnan, G., Hoteit, I., Hansen, C., Hadwiger, M., 2013. Visual analysis of uncertainties in ocean forecasts for planning and operation of off-shore structures. In: *Proceedings of IEEE Pacific Visualization Symposium 2013*, pp. 185–192.
- Hummel, M., Obermaier, H., Garth, C., Joy, K.I., 2013. Comparative visual analysis of Lagrangian transport in CFD ensembles. *IEEE Trans. Vis. Comput. Graphics* 19 (12), 2743–2752.

- Jarema, M., Demir, I., Kehrler, J., Westermann, R., 2015. Comparative visual analysis of vector field ensembles. In: Proceedings of IEEE Visual Analytics Science and Technology 2015. IEEE, pp. 81–88.
- Kehrler, J., Hauser, H., 2013. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Trans. Vis. Comput. Graphics* 19 (3), 495–513.
- Kothur, P., Sips, M., Dobslaw, H., Dransch, D., 2014. Visual analytics for comparison of ocean model output with reference data: Detecting and analyzing geophysical processes using clustering ensembles. *IEEE Trans. Vis. Comput. Graphics* 20 (12), 1893–1902.
- Kumpf, A., Rautenhaus, M., Riemer, M., Westermann, R., 2018. Visual analysis of the temporal evolution of ensemble forecast sensitivities. *IEEE Trans. Vis. Comput. Graph.* 1–11 (in press).
- Li, S., Marsaglia, N., Garth, C., Woodring, J., Clyne, J., Childs, H., 2018. Data reduction techniques for simulation, visualization and data analysis. *Comput. Graph. Forum* 37 (6), 422–447.
- Liu, R., Guo, H., Yuan, X., 2015. A bottom-up scheme for user-defined feature comparison in ensemble data. In: ACM SIGGRAPH Asia 2015 Symposium on Visualization in High Performance Computing, pp. 1–4.
- Liu, R., Guo, H., Yuan, X., 2017. User-defined feature comparison for vector field ensembles. *J. Vis.* 20 (2), 217–229.
- Liu, R., Guo, H., Zhang, J., Yuan, X., 2016. Comparative Visualization of Vector Field Ensembles Based on Longest Common Subsequence. In: Proceedings of IEEE Pacific Visualization Symposium 2016, pp. 96–103.
- Matkovic, K., Gracanin, D., Jelovic, M., Ammer, A., Lez, A., Hauser, H., 2010. Interactive visual analysis of multiple simulation runs using the simulation model view: Understanding and tuning of an electronic unit injector. *IEEE Trans. Vis. Comput. Graphics* 16 (6), 1449–1457.
- Matkovic, K., Gracanin, D., Jelovic, M., Hauser, H., 2008. Interactive visual steering-rapid visual prototyping of a common rail injection system. *IEEE Trans. Vis. Comput. Graphics* 14 (6), 1699–1706.
- Matkovic, K., Gracanin, D., Klarin, B., Hauser, H., 2009. Interactive visual analysis of complex scientific data as families of data surfaces. *IEEE Trans. Vis. Comput. Graphics* 15 (6), 1351–1358.
- Matkovic, K., Gracanin, D., Splechtna, R., Jelovic, M., Stehno, B., Hauser, H., Purgathofer, W., 2014. Visual analytics for complex engineering systems: Hybrid visual steering of simulation ensembles. *IEEE Trans. Vis. Comput. Graphics* 20 (12), 1803–1812.
- Mirzargar, M., Whitaker, R.T., Kirby, R.M., 2014. Curve boxplot: Generalization of boxplot for ensembles of curves. *IEEE Trans. Vis. Comput. Graphics* 20 (12), 2654–2663.
- Muigg, P., Kehrler, J., Oeltze, S., Piringer, H., Doleisch, H., Preim, B., Hauser, H., 2008. A four-level focus+context approach to interactive visual analysis of temporal features in large scientific data. *Comput. Graph. Forum* 27 (3), 775–782.
- Obermaier, H., Joy, K.I., 2014. Future challenges for ensemble visualization. *IEEE Comput. Graph. Appl.* 34 (3), 8–11.
- Oster, T., Lehmann, D.J., Fru, G., Theisel, H., Thévenin, D., 2014. Sparse representation and visualization for direct numerical simulation of premixed combustion. *Comput. Graph. Forum* 33 (3), 321–330.
- Poco, J., Dasgupta, A., Wei, Y., Hargrove, W., 2014a. Visual reconciliation of alternative similarity spaces in climate modeling. *IEEE Trans. Vis. Comput. Graphics* 20 (12), 1923–1932.
- Poco, J., Dasgupta, A., Wei, Y., Hargrove, W., Schwalm, C., Cook, R., Bertini, E., Silva, C., 2014b. Similarityexplorer: A visual inter-comparison tool for multifaceted climate data. *Comput. Graph. Forum* 33 (3), 341–350.
- Potter, K., Kniss, J., Riesenfeld, R., Johnson, C.R., 2010. Visualizing summary statistics and uncertainty. *Comput. Graph. Forum* 29 (3), 823–832.
- Potter, K., Wilson, A., Bremer, P.-T., Williams, D., Doutriaux, C., Pascucci, V., Johnson, C., 2009b. Visualization of uncertainty and ensemble data: Exploration of climate modeling and weather forecast data with integrated VISUS-CDAT systems. *J. Phys. Conf. Ser.* 180 (1), 1–5.
- Potter, K., Wilson, A.T., Bremer, P.-T., Williams, D.N., Doutriaux, C.M., Pascucci, V., Johnson, C.R., 2009a. Ensemble-vis: A framework for the statistical visualization of ensemble data. In: Proceedings of IEEE International Conference on Data Mining Workshops 2009, pp. 233–240.
- Rautenhaus, M., Bottinger, M., Siemen, S., Hoffman, R., Kirby, R.M., Mirzargar, M., Rober, N., Westermann, R., 2017. Visualization in meteorology – A survey of techniques and tools for data analysis tasks. *IEEE Trans. Vis. Comput. Graphics* 24 (12), 1–30.
- Ropinski, T., Oeltze, S., Preim, B., 2011. Survey of glyph-based visualization techniques for partial multivariate medical data. *Comput. Graph.* 35 (2), 392–401.
- Sanyal, J., Zhang, S., Dyer, J., Mercer, A., Amburn, P., Moorhead, R.J., 2010. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Trans. Vis. Comput. Graphics* 16 (6), 1421–1430.
- Shu, Q., Guo, H., Che, L., Yuan, X., Liu, J., Liang, J., 2016. EnsembleGraph: Interactive visual analysis of spatiotemporal behaviors in ensemble simulation data. In: Proceedings of IEEE Pacific Visualization Symposium 2016, pp. 56–63.
- Unger, A., Schumann, H., 2009. Visual support for the understanding of simulation processes. In: Proceedings of IEEE Pacific Visualization Symposium 2009, pp. 57–64.
- Wang, J., Hazarika, S., Li, C., Shen, H.-W., 2015. Visualization and visual analysis of ensemble data: A survey. *IEEE Trans. Vis. Comput. Graphics* 14 (8), 1–20.
- Wang, J., Liu, X., Shen, H.-W., Lin, G., 2017. Multi-resolution climate ensemble parameter analysis with nested parallel coordinates plots. *IEEE Trans. Vis. Comput. Graphics* 23 (1), 81–90.
- Weber, G.H., Bremer, P.-T., Pascucci, V., 2007. Topological landscapes: A terrain metaphor for scientific data. *IEEE Trans. Vis. Comput. Graphics* 13 (6), 1416–1423.
- Whitaker, R.T., Mirzargar, M., Kirby, R.M., 2013. Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. *IEEE Trans. Vis. Comput. Graphics* 19 (12), 2713–2722.
- Williams, S., Hecht, M., Petersen, M., Strelitz, R., Maltrud, M., Ahrens, J., Hlawitschka, M., Hamann, B., 2011. Visualization and analysis of eddies in a global ocean simulation. *Comput. Graph. Forum* 30 (3), 991–1000.
- Woodring, J., Ahrens, J., Figg, J., Wendelberger, J., Habib, S., Heitmann, K., 2011. In-situ sampling of a large-scale particle simulation for interactive visualization and analysis. *Comput. Graph. Forum* 30 (3), 1151–1160.
- Woodring, J., Petersen, M., Schmeißer, A., Patchett, J., Ahrens, J., Hagen, H., 2016. In situ eddy analysis in a high-resolution ocean climate model. *IEEE Trans. Vis. Comput. Graphics* 22 (1), 857–866.