Reprinted from
Journal of Imaging Science and Technology® 60(6): 060408-1–060408-11, 2016.
https://doi.org/10.2352/ISSN.2470-1173.2017.1.VDA-393
© Society for Imaging Science and Technology 2016

Effectiveness of Feature-Driven Storytelling in 3D Time-Varying Data Visualization

Li Yu, Lane Harrison, and Aidong Lu
University of North Carolina at Charlotte, Charlotte, NC 28223
E-mail: alu1@uncc.edu

Abstract. *Storytelling animation has a great potential to be widely adopted by domain scientists for exploring trends in scientific simulations. However, due to the dynamic nature and generation methods of animations, serious concerns have been raised regarding their effectiveness for analytical tasks. This has led to interactive techniques often being favored over animations, as they provide the user with complete control over the visualization. This trend in scientific visualization design has not yet considered newer algorithmic animation generation methods that are driven by the automatic analysis of data features and storytelling techniques. In this work, the authors performed an experiment which compares feature-driven storytelling animations to common interactive visualization techniques for time-varying scientific simulations. They discuss the design of the experiment, including tasks for storm-surge analysis that are representative of common scientific visualization projects. Their results illustrate the relative advantages of both feature-driven storytelling animations and interactive visualizations, which may provide useful design guidelines for future storytelling and scientific visualization techniques. © 2016 Society for Imaging Science and Technology.*

INTRODUCTION

Animation is widely used to show trends in many data-intensive applications. It also remains one of the most popular choices for end users, since it is a natural and attractive way to represent dynamic events. When studying time-varying simulation data, domain scientists often use animations to visualize temporal events. In fact, animation has become an essential part of many scientists' workflow for analyzing simulations.

Despite its widespread use with domain scientists, animation has never been a dominant or even popular approach in the scientific visualization community. Part of the reason for this trend is that animations are often simplistically generated by connecting snapshots from individual time steps, which does not meet the requirements of challenging research tasks. Another reason is that while several methods have been developed for generating animations for visualization, including user-driven editing systems[1] and automatic animation approaches,[2–4] complex animations have not been sufficiently evaluated with domain users or compared to interactive approaches.

Received July 4, 2016; accepted for publication Nov. 5, 2016; published online Dec. 7, 2016. Associate Editor: Thomas Wishgoll.

Research in visualization has also raised concerns regarding the effectiveness of animation. A number of studies have been performed to evaluate the effectiveness of animation in different research fields, such as information visualization,[5–7] graph visualization,[8–10] medical visualization,[11] hierarchical diagram differencing,[12] and visual tracking.[9] The results from these studies are mixed. While some researchers have found animation to be effective in visualizations,[5,11] others have concluded that alternate methods such as small multiples should be used for analysis tasks, regardless of the fact that animation is widespread and engaging.[6,13] The main concern is whether animation facilitates accurate perception of changes in the data[14] and supports iterative analysis.[6]

To explore the gap between the popularity of animation and concerns about its fitness for scientific analysis, we designed an experiment to evaluate the performance of feature-driven animation and interactive systems. Two systems are described in the study design, one using feature-driven storytelling animations and the other using common interaction techniques. We describe an implementation of feature-driven animation designed to satisfy general requirements for analyzing time-varying data visualization. The experiment uses two similar hurricane/storm-surge datasets and a corresponding set of tasks. The tasks are designed to represent common studies users perform in scientific visualization: examining event representation, data exploration, and reasoning about relationships between underlying data features.

Our results suggest that feature-driven storytelling animations may consistently lead to more timely results, and have comparable accuracy to interactive visualization across a variety of tasks common to scientific visualization. Given the scope of the study, we view our results as an indication that feature-driven animation can and should play a larger role in scientific visualization design, particularly as a mechanism for providing an overview and a moderate level of detail for the dominant features in time-varying scientific simulations. Although this study is based on an application of time-varying visualization (storm-surge analysis), the design of our study, including both the feature-driven animation approach and selection of visualization tasks, can be applied to other scientific visualization domains.

The long-term goal of our study on storytelling animation is to explore new visualization and interaction mechanisms that suit the changing environments of visualization applications, particularly the shorter simulation-to-analysis-to-simulation cycle of domain scientists and the continual increases in simulation size and complexity. Animation will likely remain the dominant method domain scientists use to explore time-varying simulation data. Therefore, visualization researchers should focus on studying how users work with animations and identifying where interaction is most needed in order to improve visualization design.

The article is organized as follows. We first review related work on animation, storytelling and narrative visualization, and evaluation methods in Related Work. Then, we describe our study design in Study Design and tasks for the user study in Task. We continue to present the details of the experiment, evaluation and discussion from the user study in Experiment and Discussion respectively. At the end, we conclude the work and present future work in Conclusion and Future Work.

RELATED WORK

Animation for Visualization

In scientific visualization, approaches related to animation have concentrated on creating animations and using animations to highlight data features.[3,15,16] For example, Gershon[17] presented methods for visualizing fuzzy data in an animation loop. Viola et al.[2] presented a method to focus viewpoints automatically on data features. Akiba et al.[1] presented AniViz to create animations with templates and operators. Yu et al.[4] presented an automatic animation generation approach for time-varying data visualization. Animations were also used to visualize vector datasets.[18,19]

For information visualization, animation approaches have also been explored. For example, Heer and Robertson[5] developed design guidelines for animated transitions. Lundström et al.[11] used animation to convey uncertainty in medical visualization. Blumenkrants et al.[20] created narrative visualization to study algorithms. Zongker and Salesin[21] proposed and discussed the principles and guidelines of animation for slideshow presentations. In general, given the abstract nature of data in information visualization, animation is seen as effective for presentation tasks but not for in-depth analytical tasks.

Storytelling and Narrative Visualization

The word "storytelling" has a long history and it has been introduced to visualization for improving visual communication.[22–25] While the term of narrative visualization is relatively new,[26–29] it also refers to using data stories to augment visualization as a communication medium.

Narrative structure is a central concept in both storytelling and narrative visualization. It refers to "a series of events, facts, etc., given in order and with the establishing of connections between them" from the Oxford English Dictionary and it is often simplified to structures like beginning, middle, and end in visualization systems.[26] The studies of narrative visualization have been performed from several aspects. Segel and Heer[26] investigated the design space of narrative visualization, including the genres, visual narratives, and narrative structures. Hullman and Diakopoulos[27] demonstrated visualization rhetoric as an analytical framework for understanding the effects of design techniques on end-user interpretation. Hullman et al.[28] later conducted a qualitative analysis of 42 professional narrative visualizations to gain empirical knowledge on the forms that structure and sequence took. Satyanarayan and Heer[29] developed a model of storytelling abstractions and instantiate the model in Ellipsis with a graphical interface for story authoring.

Research of storytelling and narrative visualization has also been developed and applied to applications on several fields. Yu et al.[4] generated automatic animations with narrative structures extracted from event graphs for time-varying scientific visualization. Lee et al.[30] presented a visual data storytelling process with steps involved in finding insights, turning these insights into a narrative, and communicating this narrative to an audience. Andrews and Baber[31] designed a branching comic to compare how readers recall a visual narrative. Pschetz et al.[32] developed TurningPoint to investigate narrative-driven talk planning in slideware. Spaulding and Faste used studies to prototype and build immersive design words.[33]

Evaluation of Animation and Storytelling

A number of studies have been performed to evaluate the effectiveness of animation. For example, Tversky and Morrison[14] found that animation may be ineffective for displaying events and it could be too complex and too fast to be accurately perceived. However, they did acknowledge that animation supported interactions such as close-ups, zooming, alternative perspectives and control of speed, were likely to facilitate perception and comprehension. Archambault et al.[34] performed user study on mental maps with animation and small multiples and concluded that small multiples performed significantly faster but with more errors.

In information visualization, several studies have compared animation with different visualization methods such as scatterplots[7] and dynamic graphs.[8,35] For example, Boy et al.[36] conducted experiments on the effectiveness of using "introductory stories" to engage users and found out that providing a start point of exploration with an initial story did not affect the user engagement. Heer and Robertson found that animated transitions significantly improved graphical perception.[5] Robertson et al.[6] evaluated the effectiveness of animation in trend visualization and found that small comparable visualization was the most effective approach. They discovered that animation worked well in presentation tasks but not as good as other techniques for analysis purposes.

The evaluation work of animation in scientific visualization is rare. Lundström et al.[11] employed radiologists in a study simulating the clinical task of stenosis assessment, in which they found animation technique outperformed traditional rendering in terms of assessment accuracy. Boyandin et al.[37] presented a user study analyzing findings made while exploring changes in spatial interaction with

Figure 1. The left figure shows the time step indicator—a small triangle shape. The right figure shows that animation is currently in the second of the three events marked in different colors.

flow maps using animation and small multiples. They found animation tended to enable more findings related to the geographically local events and changes between years.

Different from standard animations, our feature-driven storytelling animations are designed for interactive visualization tasks. They can be used to visualize the overall temporal trends or study the details of specific events. We believe that for data with meaningful 3D structures like scientific simulations, storytelling animation techniques can be used to visualize temporal events effectively. To the best of our knowledge, no significant work has been done to evaluate the efficacy and effectiveness of animation in scientific visualization.

STUDY DESIGN

To provide the necessary background of our study, we start with a brief introduction of the storytelling animation technique and application domain used in the study.

Storytelling Guidelines for Animation

For each task, an animation is generated based on the associated events and features. All the animations used in the study are generated by following three design guidelines:

- Start with an overview of entire duration
- End with a focused view in relevant time duration
- Include all relevant data attributes in at least one segment of the animation sequence.

These guidelines are consistent with the procedure of data exploration that is often followed in the visualization community. They provide a storytelling feature to the animations by introducing an event first and visualizing additional details gradually.

We choose a typical application of scientific visualization, hurricane, and storm-surge simulation, to study the effects of storytelling animation. We follow the method used in Ref. 4 to generate storytelling animations, which can be applied to a wide range of hurricane and storm-surge simulations. This approach characterizes temporal patterns from different variables and scales in a time-varying sequence to create an event graph. Based on the animation design guidelines described above, the event graph is traversed to determine the sequence of events shown in the animation. Animations are generated automatically with changing views and smooth transitions. As shown in Figure 1, a time indicator is provided to show the current time step number and a sequence duration bar is provided at the bottom of the panel to show the remaining time in the current event. Details of animations for each task in the experiment are provided in Task.

Specifically, the animation system contains two panels: a 3D rendering panel and a temporal trend panel, which are the same as the interactive visualization system. The 3D rendering panel adopts the same rendering scheme for all the data attributes and time steps, in the sense that all the rendering effects are exactly the same. The main difference is that for the animation system, the 3D rendering panel displays animation, while for the interactive system, the 3D rendering panel displays the rendering from a selected time step. The control panel is disabled for the animation system to avoid the confusion of different animation effects.

Datasets

Two storm-surge simulation datasets are used in the experiment: Hurricane Isabel (2003) and Hurricane Irene (2011). Several features of these two datasets make them suitable for our study.

First, the two simulation datasets are similar. Both datasets cover the same region, whose terrain model consists of more than 260,000 vertexes and 520,000 triangles. They also contain the same set of four key variables: elevation, wind vectors, atmospheric pressure, and depth-average velocity. The number of time steps from the two datasets are also comparable; Isabel contains 396 time steps and Irene contains 336. Based on these observations, we expect that the choice of datasets did not affect the performance of participants in the user study.

Second, because the two hurricanes not exactly alike, we can design a set of comparable tasks with different answers for each dataset. For example, the eye paths of the two hurricanes differ completely. This means that the participants will not be able to answer the questions for one dataset by using the information they acquire from the other dataset, ensuring that the order of the datasets does not affect the results of the experiment.

Data Features

To generate animations for the evaluation, the relevant data features need to be automatically modeled and computed. As shown in Figure 2, storm-surge simulations provide complex 3D scenarios which simulate the behavior of water elevation given different hurricane-related parameters. We describe several data features that are important to our study in the following.

Back Surge Modeling

Storm surge is an abnormal rise of water generated by a storm, over and above the predicted astronomical tide. Near the coast, people are especially concerned about "back surge," which drives considerable amount of water back to the river from the ocean and may cause serious flooding. Simply speaking, the back surge travels in the opposite direction of

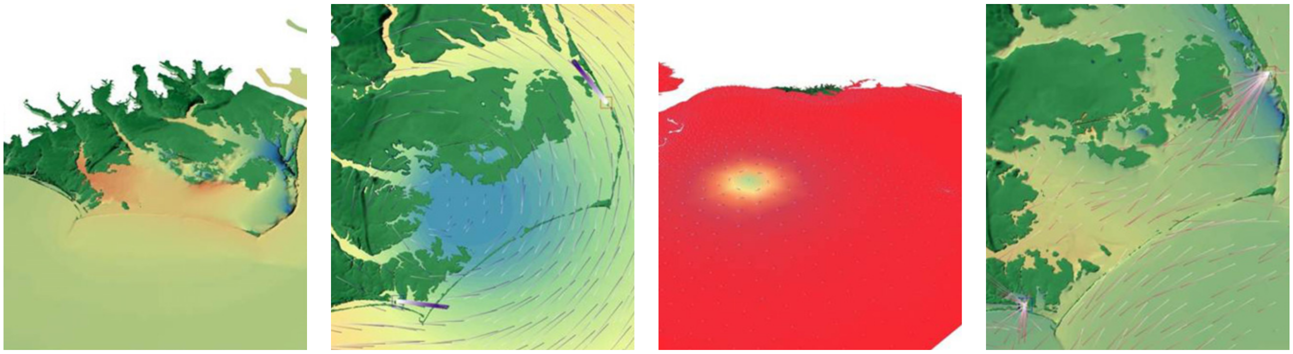


Figure 2. Example visualizations of elevation, wind vector, atmospheric pressure, and velocity parameters.

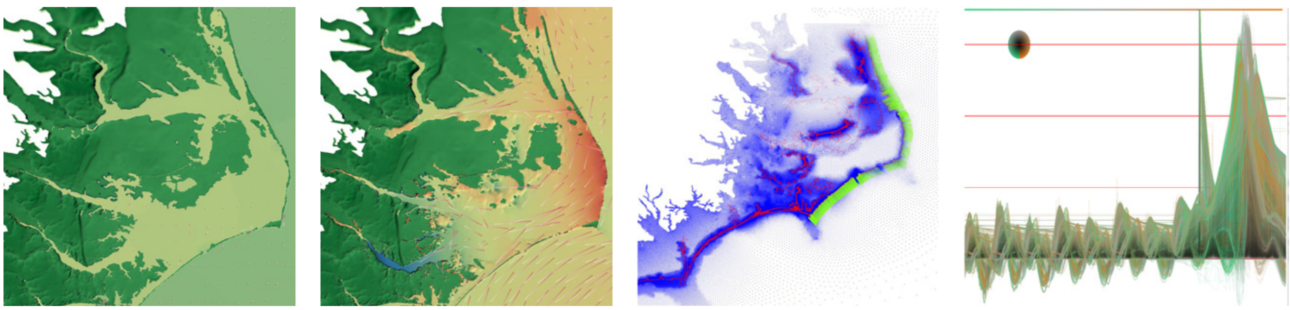


Figure 3. The two images on the left show example renderings of normal and back surge conditions. The two images on the right provide the modeling of back surge—one shows the vertexes in the grid and the other shows the elevation changes and tide directions.

normal surges. Back surge is an event that appears frequently near the North Carolina coast, specifically near the Outer Banks. It can generally be found after the eye has passed and the overall water velocity is directly perpendicular to the hurricane's direction. Figure 3 provides examples of the Outer Banks under normal and back surge conditions. We use the temporal features of water velocity and direction to model back surge with the following three steps.

First, we extract the boundary of Outer Banks by selecting vertexes whose heights are close to sea surface elevation. The water velocity directions at the boundary are used to indicate the direction of surge. In Fig. 3 (the third image), the grid of storm-surge model is shown. Second, the direction of the tides is simulated with depth-average velocity variable. We use these two angles to determine whether the tides are traveling into land or going out from land. If the angle between tide direction and the Outer Banks outgoing direction is smaller than 90°, we consider it to be water coming out of the land as back surge; otherwise, water is coming toward land. Fig. 3 (the fourth image) visualizes the temporal changes of elevation values on all the boundaries. Third, we accumulate the depth-average velocity by treating outgoing directions positive and ingoing directions negative. The back surge can be identified at the local maximum points beyond the range of normal water elevations.

Inundation

Inundation is another important feature in the simulation, as it is crucial to emergency response and evacuation planning.

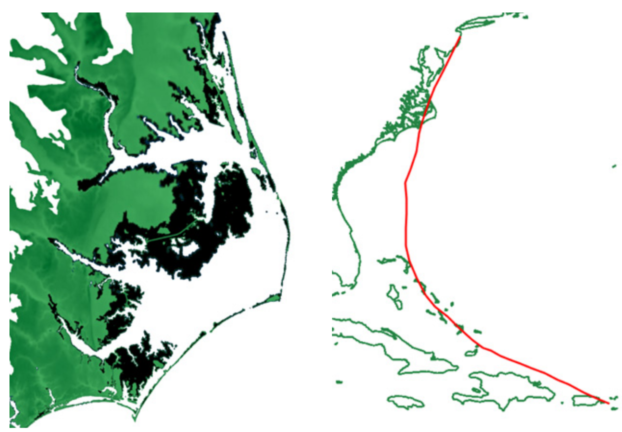


Figure 4. The inundation area of Isabel (left) and the path of Irene extracted from simulations (right).

We compute the inundation area by comparing the water elevation height of a vertex with the height of land. Figure 4 (left) shows the inundation area caused by Hurricane Isabel.

Hurricane path

The hurricane path can be computed by connecting the centers of hurricane eye from all the time steps. We identify the hurricane eye through the maximum wind velocity attribute and smooth the path using a Gaussian function. Based on the hurricane path we extracted, we further calculate the speed and locations of hurricane eye along the path. Fig. 4 (right) shows the extracted path of Irene.

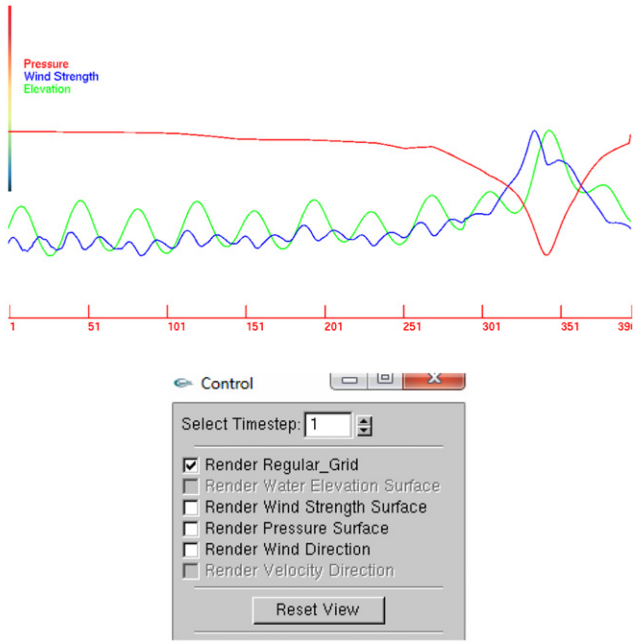


Figure 5. System interface—temporal trend panel shown at the top and the control panel shown at the bottom.

Control System: Interactive Visualization

In order to evaluate how users can effectively analyze data with animations, it is important to describe the supported interaction techniques and animations and to ensure that differences in the two compared methods are understood with regard to the outcome of the study.

The interactive visualization is selected to compare with animations as it is the most commonly used approach to study time-varying scientific simulations. The design of the interactive visualization system also adopts typical visualization systems with three panels: a 3D rendering panel, a temporal trend panel, and a control panel. The last two panels are shown in Figure 5.

Rendering Panel The 3D rendering panel visualizes all the involved data attributes from a selected time step. Standard interactions are provided, including rotation, zooming in/out, viewpoint selection, selection of time step, and selection of data attribute such as the examples in Fig. 2. We only provide these standard interactions to direct participants to focus on the tasks. We also provide a time indicator to specify the selected time step.

Temporal Trend Panel The temporal trend panel presents 2D curves of data fluctuations. This panel is included in the system to plot different variables of storm-surge simulations. The temporal trend curves are updated automatically according to the tasks. For example, for the task of visualizing surges, the curve of average water elevation is displayed in the temporal trend panel.

Control Panel The control panel is for selecting data attributes, time steps, and reset 3D view. Our interactive visualization system only enables relevant data attributes for each task to direct the focus of participants on the same set of data attributes.

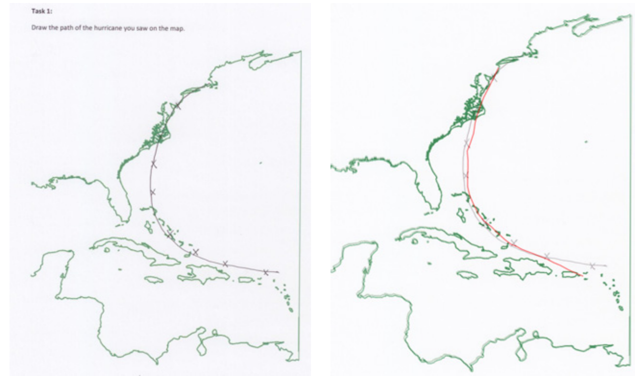


Figure 6. Task 1—Left: An example of participant's answer on a map of the Outer Banks. Right: Overlapped image of ground truth image with participant's answer. This image was then used to compute a pixel difference between the correct and participant's answers.

TASKS

Our user study consists of three categories of tasks, including representation, exploration, and reasoning tasks, which are problems that users frequently encounter in scientific visualization. Each category contains two tasks. The categories are designed and presented in increasing difficulty during the user study. In the following, we describe each task in detail and how the storytelling animations are generated to support the tasks.

Category 1—Representation

One important usage of visualization is to facilitate the accurate representation (and subsequent perception) of features in the data.[6] The purpose of these tasks is to examine how well participants understand data features from the storytelling animation or interactive visualization.

Task 1: The instruction to the participants for this task is: "Use the interactive visualization system or animation to visualize the hurricane, and draw the path of the hurricane on the map manually." As shown in Figure 6, the subjects are provided with a map on an A4 paper. They can mark locations of the hurricane eye at different time steps and connect them as a line. The participant's answer is further compared with the ground truth shown as the red line for evaluating accuracy.

The storytelling animation was designed to use the temporal trend panel to show two main attributes: atmospheric pressure and wind strength, as they were the most relevant factors related to this task. The animation consisted of two phases. Each phase displayed the hurricane wind strength and atmospheric pressure with a bird's-eye view of the entire grid of the entire time duration respectively.

Task 2: The instruction to the participants is: "Use the interactive visualization system or animation to examine the inundation areas, and mark all the inundation areas within a blue circle."

The storytelling animation was designed to use the control panel to show the water elevation, as it was the only attribute related to this task. The average of elevation of every time step was shown in the temporal trend panel. The animation of inundation along North Carolina was shown in three phases. We start with an overview of the North Carolina coast through the entire time duration. The second phase showed a more focused view on areas around Outer Banks: the Pamlico and Albemarle Sound. The last phase displayed a view on North Carolina coast after the surges. The participants were asked to report the storm-surge time duration and where the highest possibilities of inundations were.

Category 2—Exploration

This category focuses on analyzing data with storytelling animation and interactive visualization. Participants are asked to explore the data to determine whether certain features or events existed in the simulation. In addition, the participants are asked to explore the relationships between different variables, such as wind vectors and elevations.

Task 3: The instruction to the participants for this task is: "Did you find back surge during the storm surge? If yes, write down the starting time step and ending time step of the back surge, and mark where back surge appeared on the map." In this task, the participants were given the definition of back surge and the average value of elevation of every time step was displayed. Participants were asked to find specific time steps and locations of the beginning and ending of back surges. Two sample answers from participants were shown in Figure 8.

The storytelling animation was designed to contain three phases: the first showed an overview of the North Carolina coast for the entire time duration; the second phase used a focused view around the back surge area detected by our algorithm and displayed velocity vectors along with elevation surfaces; the last phase showed the elevation changes after the hurricane passed through the coastal area.

Task 4: The instruction to the participants for this task is: "Describe the relationships (location, strength, and height) between the hurricane eye and the highest elevation."

In this task, the participants were asked to identify and explain relationships between wind vectors and water elevations. The wind vectors clearly showed the hurricane eye and eye wall and the elevation surface showed when and where the ocean level changed. Two sample answers from the participants of this task were shown in Figure 9. Temporal curves of elevation and atmospheric pressure changes were displayed as hints to the participants. The interactive system enabled participants to explore elevation surface and wind vectors.

The storytelling animation was designed to use two phases to describe this event. First, the elevation and wind direction were rendered using a focused view on North Carolina coast. The second was a dynamic view following the hurricane eye to observe the wind in the second phase.

Category 3—Reasoning

In this category, the participants examined low level details to determine various relationships for certain locations in the simulation. We focused on if and how fast the participants could identify reasons for different patterns of two locations during the hurricane. The overall changes of relevant variables in the detailed area were also shown to the participants.

Task 5: The two locations were NOAA water level observation stations on the Outer Banks. One is Oregon Marina Inlet, the other is Beaufort. The vertexes around these two areas were extracted and the overall changes of water level were shown to the participants. An overview of how the hurricane traveled was presented to the participants and then the relevant variables during the storm-surge time were displayed. The instruction to the participants for this task is: "Why are the elevation and pressure changes different from each other for those two locations?"

Figure 10 showed the locations of observation stations on Outer Banks and the temporal trends of two different attributes. The storytelling animation was designed to display a focused view of these two locations rendering relevant variables with the locations highlighted. The animation duration included overall changes followed by changes during the storm-surge time period using three phases. The first phase displayed the overall changes over the entire time period. The second phase showed the elevation with velocity vectors to help participants find the reason for the differences in the elevation curves. The last phase displayed the wind vectors with atmospheric pressure to help participant inspect the pressure changes.

Task 6: There are two canals in North Carolina which could be dramatically affected by storm surge. One was the Pungo River Canal; the other was Adam Creek Canal. The two canals were impacted differently in the hurricane datasets because of their relative locations to the hurricane path. The vertexes around these two areas were extracted and the overall changes of these vertexes were shown to the participants. The participants were asked to explain the reason why the two canals were affected differently during the storm surge. The instruction to the participants for this task is: "Why does the Pungo River Canal and Adam Creek Canal have different changes in elevation?"

Figure 11 showed the locations of the two canals in North Carolina and the temporal trends of two different attributes. The storytelling animation was designed to show an overview of changes during the entire time duration, changes of elevation surfaces, and elevation and water velocity of each location during the surge time duration with a closer view.

EXPERIMENT

Using the previously described tasks, we conducted an in-lab controlled experiment to compare our two visualization systems for time-varying data visualization: feature-driven animated visualization and interactive visualization.

Participants

Participants consisted of undergraduate and graduate students from the Computer Science department. There were

Table I. Average time (seconds) spent on one task and accuracy of animation (A) and interaction (I) systems of each participant (P).

P	A Time	I Time	A Acc	I Acc
1	114.5	340.17	52.44	58.15
2	86.67	144.67	44.55	30.38
3	91.17	276.17	38.26	32.56
4	224.33	185.83	52.12	74.42
5	189.33	422	42.01	55.26
6	160.33	214.33	59.29	51.53
7	130.33	221.33	51.07	38.04
8	311.67	237.17	51.08	69.47
9	124.5	301.33	57.33	56.35
10	266.17	172	90.41	79.62
11	185.83	337.83	81.26	55.01
12	167	191.5	53.44	65.23

12 participants (10 male, 2 female) with an average age of 26.5 and standard deviation of 2.94 (maximum 29, minimum 19). Estimated total participation time (training, tasks, and debriefing) was 45 minutes. During the tasks, participants were asked to work as quickly and accurately as possible. However, due to the complexity of the tasks, no strict time limit was given.

Procedure

A training session was provided before the tasks. In training, example snapshots of each variable and a demo animation were shown. Each participant was also given several minutes to learn the interactions for the animation and control systems.

Each of the six tasks contained two sections: one with the animation and the other with the interactive system. The order of the two datasets was randomly chosen at the beginning of the participant's session, and it was alternated between interaction/animation sections. Furthermore, the order of systems presented (animated/interactive) was randomly chosen. Although no task was dependent on another, each participant completed them in the same order (tasks 1–6) for consistency.

The same procedure was used for administering each task. First, participants were given time to read the task and questions were encouraged before starting the task.

During the animation section, participants watched the animation repeatedly (with pause, fast-forward, and rewind controls) until they completed the task. For the interactive section, participants utilized the provided controls to explore different views and variables in order to complete the tasks. To better compare the systems, data variables that were not directly relevant to the task were disabled in the interactive sections.

All experiments were conducted in the same room and on the same machine (resolution 1920 × 1280). Upon completing a section, participants were given answer sheets to record their results before moving to the next task.

Table II. Task completion time (in seconds) and standard error.

Task	A (M)	I (M)	A (SD)	I (SD)	p
All	170.99	257.36	102.21	127.62	$5.04E^{-06}$
1	94.75	190.92	55.57	63.14	0.0025
2	171.17	238.5	124.32	112.16	0.1056
3	162.83	218.92	80.84	110.34	0.1058
4	168.67	279.83	135.23	154.14	0.0489
5	222.83	321.58	91.01	152.64	0.0407
6	205.67	294.42	71.52	126.33	0.0095

Results

Our study adheres to a within-subjects design, since participants completed the same tasks with both the animation and interactive systems. We use paired t-test[38] throughout this section to compare values across tasks, as it is commonly used to compare two population means when they are correlated (the samples are the matched pairs in our study).

The average data for all participants is included in Table I. We divide our analysis into two parts: completion time and accuracy. The implications of the results are explored further in the discussion.

Completion Time

This study includes 144 completion times, 2 for each of the 6 tasks (one for animation, one for interaction) and 12 participants. A strong significant effect is found for completion time with p-value $5.04E^{-06}$, indicating that animation tasks are being completed more quickly than interaction tasks ($M = 170.99$, $SD = 102.21$ and $M = 257.36$, $SD = 127.62$, respectively). This result supports our hypothesis that animation supports more timely performance than interactive visualizations.

Table II shows the average completion time, standard deviation, and t-test results of each task. The t-test results show significant differences in overall completion time and for tasks 1, 4, 5, and 6. While the average completion times for tasks 2 and 3 are still lower for animation than the interactive system, the differences are not significant (see Table II). Figure 12 shows the average completion times and standard error for each of the tasks.

The average completion time for tasks using the Isabel dataset is 3 minutes and 31 seconds, while for Irene it was 3 minutes and 39 seconds. The closeness of these values indicates that our results were not affected by using different datasets.

The average completion time for each task category (representation, exploration, and reasoning) is 176, 207, and 262 seconds, respectively. These results are consistent with the difficulty of each category.

Accuracy

We describe how we scored each task before presenting the analysis of task accuracy.

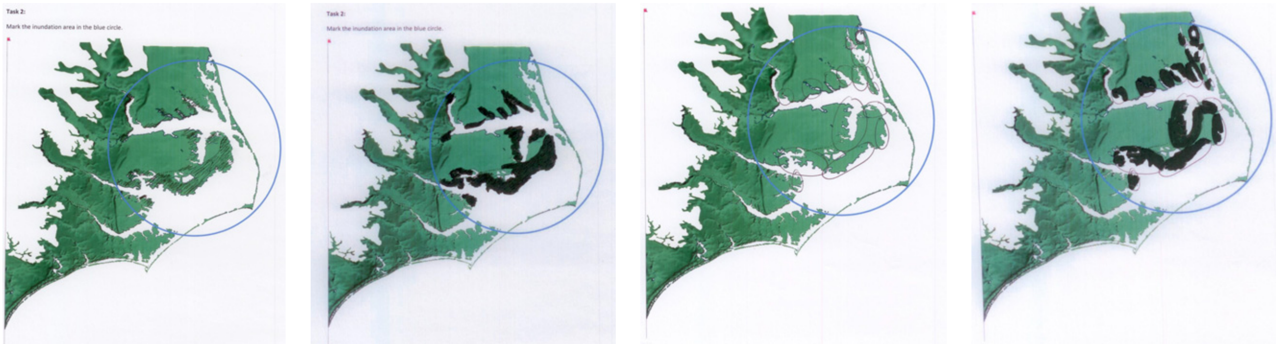


Figure 7. Task 2—two examples of participants' answers. Among each pair, the left image shows the answer of the participant, and the right image is our processed result with the participant answers painted black. These images were then compared with the ground truth image.

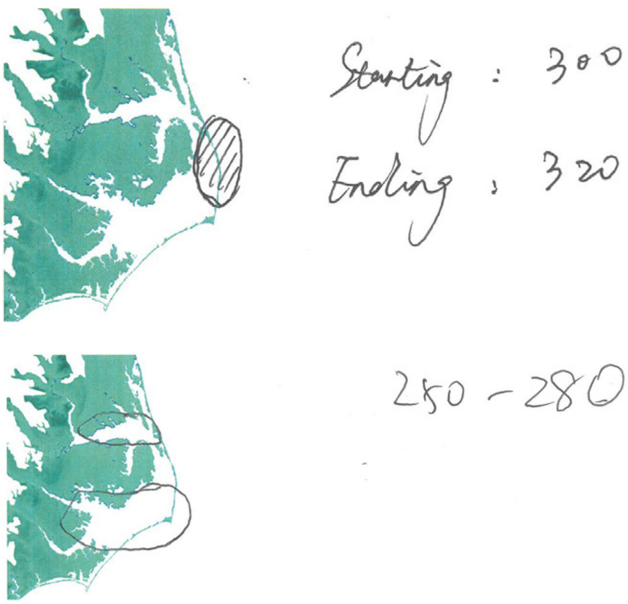


Figure 8. Task 3—example answers from participants. Participants marked back surge area and wrote down the time duration.

Representation (1 & 2): Since both tasks in this category require participants to draw answers manually, we use the following pixel-based method to compare their answers with the ground truth. For task 1, 20 points are evenly sampled along the path of ground truth. Then, for each of the 20 points, we search for the closest point from the participant's drawing. The sum of the distances between all the point pairs gives a quantitative result for task 1. For task 2, we compare the areas (number of pixels) from participant answers and the ground truth. The ground truth and a sample answer for task 2 is shown in Figure 7.

Exploration (3 & 4): The grading for task 3 is achieved as follows: for each of the answers on the starting time step, the ending time step, and location, credits are weighted as 33%. We allowed a ±5 time step differences in answering time step values. The answers for task 4 were that there were several different phases of the relationship between the highest elevation and hurricane, such as highest elevation

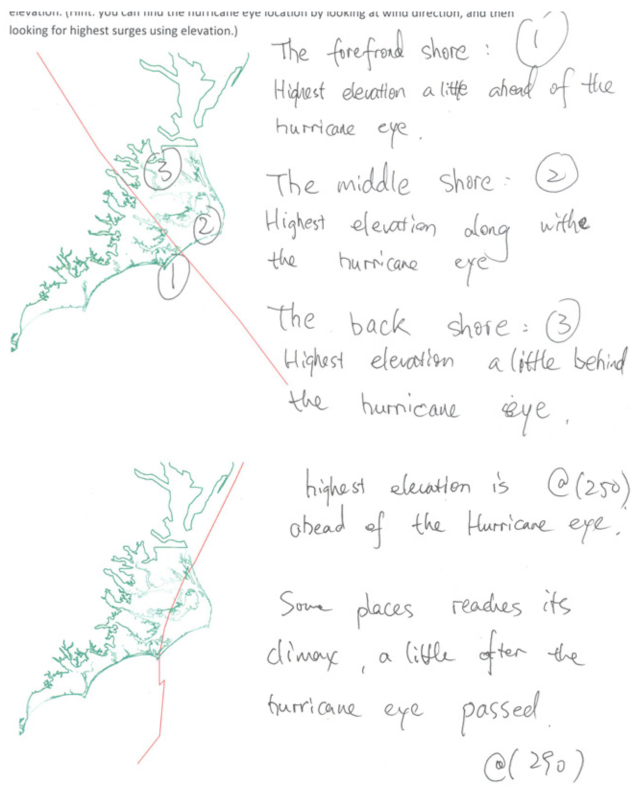
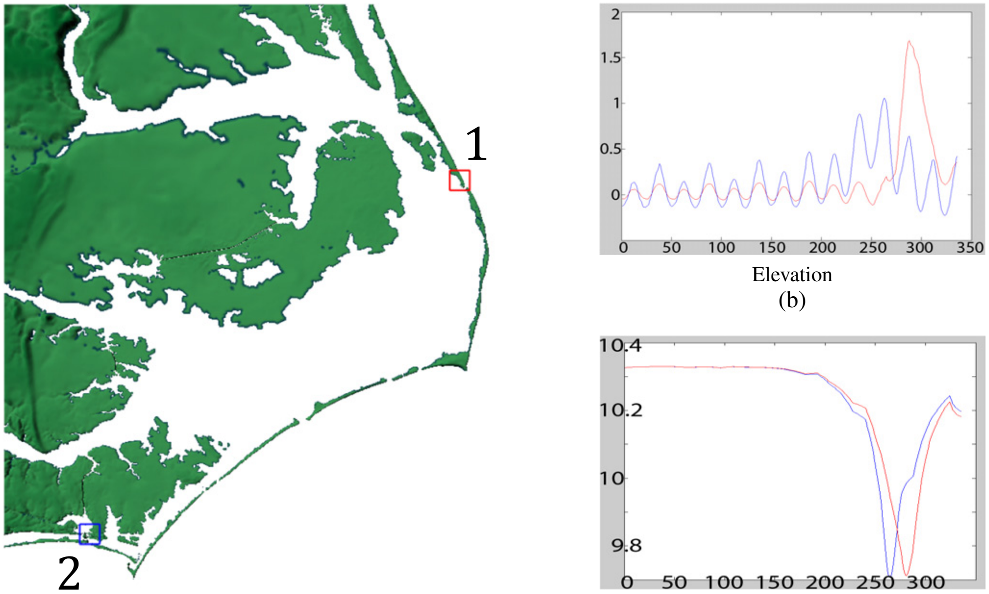


Figure 9. Task 4—Example answers from participants. Participants draw a red line to indicate the hurricane path and wrote down the answers. The first answer was: "The fore front shore: Highest elevation a little ahead of the hurricane eye. The middle shore: Highest elevation along with the hurricane eye. The back shore: Highest elevation a little behind the hurricane eye." The second answer was: "Highest elevation is at 250 ahead of the hurricane eye. Some places reduces its climax, a little after the hurricane eye passed at 290."

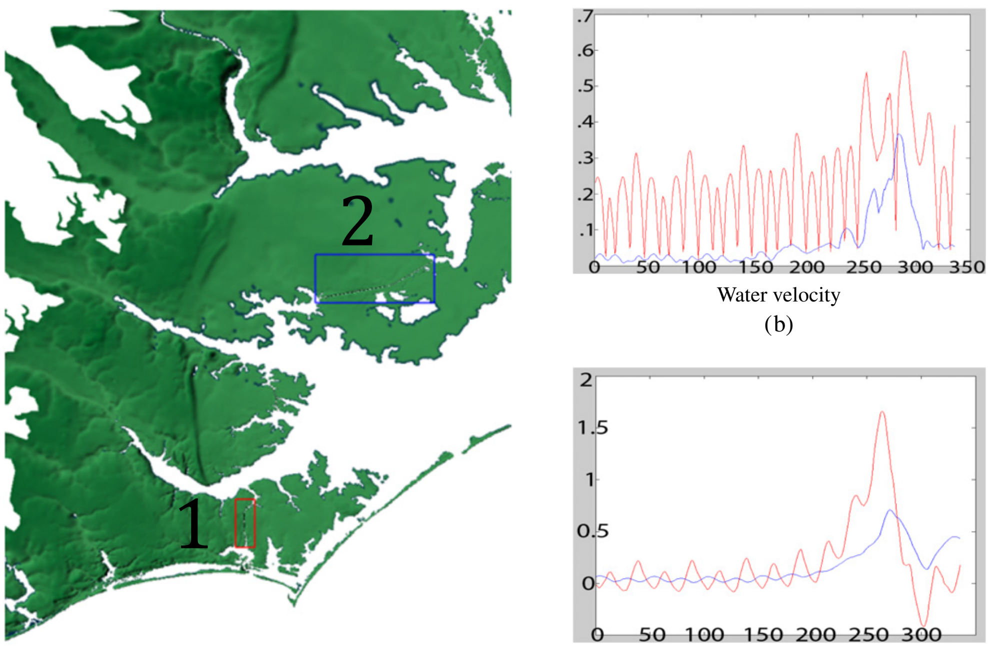
appeared both before and after the hurricane eye along the path. We gave partial credits to each correct description.

Reasoning (5 & 6): Both tasks 5 and 6 require participants to describe differences in the apparent impact of the hurricane on variables such as water elevation or pressure in two areas. Correct answers described the relative position of the hurricane eye, as well as the time when the hurricane passed.



Task 5 Elevation Atmospheric pressure
(a) (b) (c)

Figure 10. Task 5: (a) The two locations in task 5 on the Outer Banks, marked with numbers and in red or blue box. (b) and (c) show the temporal trends of elevation changes and atmospheric pressure of these two locations.



Task 6 Water velocity Elevation
(a) (b) (c)

Figure 11. Task 6: (a) The locations of the two canals in task 6, marked with numbers and in red or blue box. (b) and (c) show the temporal trends of water velocities and elevation changes of these two locations.

Task 5 required the participants to figure out two features of the hurricanes. The first feature is that one storm passed both locations. The second feature is that the hurricane eye is closer to one location than the other which causes the effects of different pressure. For scoring, 50 points were removed for each error.

Task 6 asked for two features: one that one storm created big splats on one canal and the other is that two canals are laying at different directions. Each of the reasons were credited 50 points.

The accuracy study includes 144 scores, 2 for each of the 6 tasks (one for animation, one for interaction) and

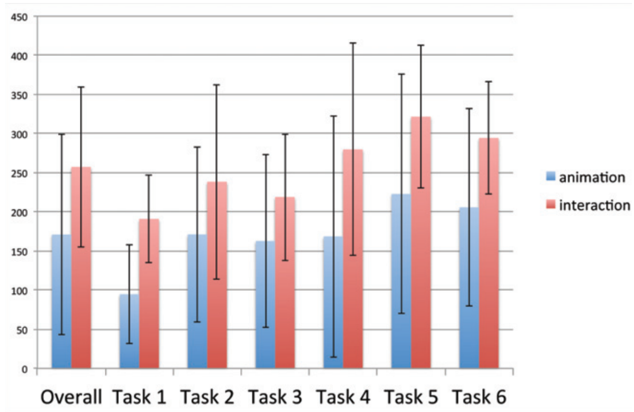


Figure 12. Task completion time and standard errors.

Table III. Task accuracy and standard deviation.

Task	A (M)	I (M)	A (SD)	I (SD)	*p*
All	56.10	55.50	34.01	34.96	0.43
1	80.07	84.07	17.61	17.68	0.3381
2	63.97	61.86	9.40	6.97	0.3025
3	60.75	52.5	29.94	34.58	0.2479
4	52.67	51.25	39.00	41.62	0.3961
5	58.33	62.5	27.64	29.76	0.3371
6	20.83	20.83	37.96	32.00	0.5

12 participants. No significant effect was found for overall accuracy (p-value 0.43) with animation and interaction having similar scores ($M = 56.10$, $SD = 34.01$ and $M = 55.50$, $SD = 34.96$, respectively). This result does not support our hypothesis that interactive visualization supports more accurate performance than animated visualizations.

Table III shows the average accuracy, standard deviation, and t-test results of each task. The t-test results show no significant differences at neither the overall nor the task level. Figure 13 shows the average accuracy scores and standard error bars for each of the tasks.

The average accuracy for tasks using the Isabel dataset was 52.87%, while for Irene it was 58.74%. Similar to the completion time, the accuracy results indicate that our study was not affected by using different datasets.

The average accuracy for each task category (representation, exploration, and reasoning) was 72.49%, 54.29%, and 40.63%, respectively. These results are also consistent with the difficulty of each category.

DISCUSSION

This section reviews the results of our experiment and touches on some fundamental questions about the role of animation and interaction in visualization.

We hypothesized that interaction would lead to more accurate results on Exploration and Reasoning tasks, and the feature-driven animation would lead to more timely results. While we did find that feature-driven animation consistently

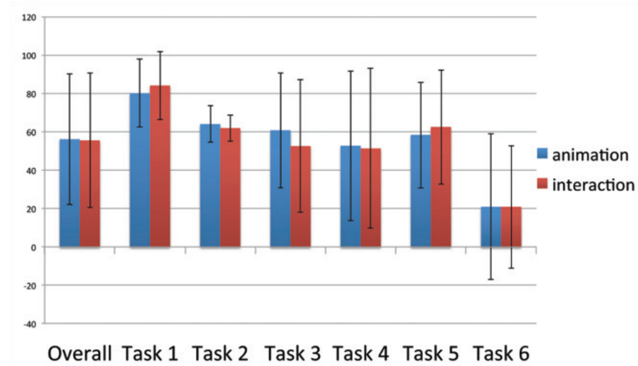


Figure 13. Task accuracy and standard error bars.

led to faster results ($p < 0.001$), the accuracy between the animation and interactive systems was comparable ($p > 0.05$).

The similarity in accuracy does not indicate that animation can simply be substituted for interaction. Instead, it indicates that for several tasks in which interaction is commonly used, feature-driven animation may support the user equally well. In addition, the experiment required the exploration and reasoning tasks to be well defined, and it is likely that tasks which are less defined will benefit from interactive capabilities.

Tasks using the animation system were completed faster than those using the interactive system. This effect held regardless of whether animation or interaction was first used on the task, since the order of the systems presented was random. Since accuracy between the systems was comparable, this supports our contention that feature-driven animation can and should play a larger role in the design of scientific visualizations.

Similar to the previous studies of animation, users often described animation as fun and engaging. This is important to visualization for directing user attention and designing effective analysis approaches. Particularly for time-varying data research, where temporal changes are pervasive, animation remains a natural way to represent and analyze the characteristics of data changes across time.

Identifying suitable ways to generate animations and incorporating animation into the interactive exploration process should prove valuable for time-varying data visualization. While many feature-driven approaches for scientific visualization have been developed, the problem of how to incorporate domain knowledge and advanced computing models is still open. Finally, beyond our basic design of feature-driven animation, additional storytelling techniques should be studied to explore methods for generating animations that is yet more intuitive for users to interpret.

CONCLUSION AND FUTURE WORK

This article presents an experiment to compare feature-driven storytelling animations to interactive visualization for studying time-varying 3D simulations. Two systems are compared using three categories of visualization tasks, including simple representation, exploration, and reasoning tasks. The results of experiment show that feature-driven animations consistently led to more timely results with comparable accuracy to the interactive system. Since interactive visualization has been the dominant approach used in scientific visualization, the results highlight the promise of feature-driven animation for future design.

In the future, we plan to conduct an experiment with students and faculty from our Meteorology Department to explore the differences in animation and interactive visualization for participants who have varying levels of domain knowledge. We believe that animation, as one of the most popular tools for scientists working with simulation data, should be studied and improved in order to become a more effective part of time-varying visualization.

ACKNOWLEDGMENTS

The authors thank the reviewers for their helpful comments. This material is based upon work supported by DHS Center of Excellence—Natural Disasters, Coastal Infrastructure and Emergency Management (DIEM).

REFERENCES

[1] H. Akiba, C. Wang, and K.-L. Ma, "Aniviz: A template-based animation tool for volume visualization," IEEE Comput. Graph. Appl. **99** (2009).

[2] I. Viola, M. Feixas, M. Sbert, and M. E. Gröller, "Importance-driven focus of attention," *Proc. IEEE Visualization* (IEEE, Piscataway, NJ, 2006), pp. 933–940.

[3] J. Woodring and H.-W. Shen, "Incorporating highlighting animations into static visualizations," Proc. SPIE **6495** (2007).

[4] L. Yu, A. Lu, W. Ribarsky, and W. Chen, "Automatic animation for time-varying data visualization," Comput. Graph. Forum **29**, 2271–2280 (2010).

[5] J. Heer and G. Robertson, "Animated transitions in statistical data graphics," IEEE Trans. Vis. Comput. Graphics **13**, 1240–1247 (2007).

[6] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko, "Effectiveness of animation in trend visualization," IEEE Trans. Vis. Comput. Graphics **14**, 1325–1332 (2008).

[7] N. Elmqvist, P. Dragicevic, and J.-D. Fekete, "Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation," IEEE Trans. Vis. Comput. Graphics **14**, 1539–1148 (2008).

[8] M. Farrugia and A. Quigley, "Effective temporal graph layout: a comparative study of animation versus static display methods," J. Inf. Vis. **10**, 47–64 (2011).

[9] F. Chevalier, P. Dragicevic, and S. Franconeri, "The not-so-staggering effect of staggered animated transitions on visual tracking," IEEE Trans. Vis. Comput. Graphics **20**, 2241–2250 (2014).

[10] B. Bach, J.-D. Fekete, and E. Pietriga, "Graphdiaries: Animated transitions and temporal navigation for dynamic networks," IEEE Trans. Vis. Comput. Graphics **20**, 740–754 (2014).

[11] C. Lundström, P. Ljung, A. Persson, and A. Ynnerman, "Uncertainty visualization in medical volume rendering using probabilistic animation," IEEE Trans. Vis. Comput. Graphics **13**, 1648–1655 (2007).

[12] L. Zaman, A. Kalra, and W. Stuerzlinger, "The effect of animation, dual view, difference layers, and relative re-layout in hierarchical diagram differencing," *Proc. Graphics Interface 2011, GI '11* (Canada, 2011), pp. 183–190.

[13] C. A. Blok, "Interactive animation to visually explore time series of satellite imagery," *Proc. 8th Int'l. Conf. Visual Information and Information Systems*, VISUAL'05 (Amsterdam, 2006), pp. 71–82.

[14] B. Tversky, J. B. Morrison, and M. Betrancourt, "Animation: can it facilitate?," Int. J. Hum.-Comput. Stud. **57**, 247–262 (2002).

[15] E. B. Lum, A. Stompel, and K. L. Ma, "Kinetic visualization: a technique for illustrating 3d shape and structure," *VIS '02: Proc. Conf. on Visualization '02* (Boston, 2002), pp. 435–442.

[16] C. D. Correa and D. Silver, "Dataset traversal with motion-controlled transfer functions," IEEE Vis. Conf. **0**, 46 (2005).

[17] N. D. Gershon, "Visualization of fuzzy data using generalized animation," *VIS '92: Proc. 3rd Conf. on Visualization '92* (Boston, 1992), pp. 268–273.

[18] W. Lefer, B. Jobard, and C. Leduc, "High-quality animation of 2d steady vector fields," IEEE Trans. Vis. Comput. Graphics **10**, 2–14 (2004).

[19] S. Bachthaler and D. Weiskopf, "Animation of orthogonal texture patterns for vector field visualization," IEEE Tran. Vis. Comput. Graphics **14**, 741–755 (2008).

[20] M. Blumenkrants, H. Starovisky, and A. Shamir, "Narrative algorithm visualization," *SoftVis '06: Proc. 2006 ACM Symposium on Software Visualization* (Brighton, 2006).

[21] D. Zongker, D. Zongker, M. Agrawala, B. Curless, D. H. Salesin, D. H. Salesin, P. David, and H. Salesin, "Creating animation for presentations," *ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (San Diego, 2003), pp. 298–308.

[22] M. Riedl and R. Young, "From linear story generation to branching story graphs," IEEE Comput. Graph. Appl. **26**, 23–31 (2006).

[23] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis, "Scientific storytelling using visualization," IEEE Comput. Graph. Appl. **32**, 12–19 (2012).

[24] R. Kosara and J. Mackinlay, "Storytelling: The next step for visualization," Computer **46**, 44–50 (2013).

[25] B. Lee, N. Riche, P. Isenberg, and S. Carpendale, "More than telling a story: Transforming data into visually shared stories," IEEE Comput. Graph. Appl. **35**, 84–90 (2015).

[26] E. Segel and J. Heer, "Narrative visualization: Telling stories with data," IEEE Trans. Vis. Comput. Graphics **16**, 1139–1148 (2010).

[27] J. Hullman and N. Diakopoulos, "Visualization rhetoric: Framing effects in narrative visualization," IEEE Trans. Vis. Comput. Graphics **17**, 2231–2240 (2011).

[28] J. Hullman, S. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar, "A deeper understanding of sequence in narrative visualization," IEEE Trans. Vis. Comput. Graphics **19**, 2406–2415 (2013).

[29] A. Satyanarayan and J. Heer, "Authoring narrative visualizations with ellipsis," Comput. Graph. Forum **33**, 361–370 (2014).

[30] B. Lee, R. H. Kazi, and G. Smith, "Sketchstory: Telling more engaging stories with data through freeform sketching," IEEE Trans. Vis. Comput. Graphics **19**, 2416–2425 (2013).

[31] D. Andrews and C. Baber, "Visualizing interactive narratives: employing a branching comic to tell a story and show its readings," *Proc. 32nd Annual ACM Conf. on Human Factors in Computing Systems* (ACM, New York, NY, 2014), pp. 1895–1904.

[32] L. Pschetz, K. Yatani, and D. Edge, "Turningpoint: narrative-driven presentation planning," *Proc. 32nd Annual ACM Conf. Human Factors in Computing Systems* (ACM, New York, NY, 2014), pp. 1591–1594.

[33] E. Spaulding and H. Faste, "Design-driven narrative: using stories to prototype and build immersive design worlds," *Proc. SIGCHI Conf. on Human Factors in Computing Systems* (ACM, New York, NY, 2013), pp. 2843–2852.

[34] D. Archambault, H. Purchase, and B. Pinaud, "Animation, small multiples, and the effect of mental map preservation in dynamic graphs," IEEE Trans. Vis. Comput. Graphics **17**, 539–552 (2011).

[35] C. Erten, P. J. Harding, S. G. Kobourov, K. Wampler, and G. Yee, "Graphael: Graph animations with evolving layouts," *11th Symposium on Graph Drawing* (2003), pp. 98–110.

[36] J. Boy, J.-D. Fekete, and F. Detienne, "Storytelling in information visualizations: Does it engage users to explore data?," *Proc. 33rd ACM Conf. on Human Factors in Computing Systems (CHI 2015)* (ACM, New York, NY, 2015).

[37] I. Boyandin, E. Bertini, and D. Lalanne, "A qualitative study on the exploration of temporal changes in flow maps with animation and small-multiples," Comput. Graph. Forum **31**, 1005–1014 (2012).

[38] M. O'Mahony, *Sensory Evaluation of Food: Statistical Methods and Procedures* (CRC Press, New York, NY, 1986).