J Vis (2020) 23:707–720
https://doi.org/10.1007/s12650-020-00644-z

REGULAR PAPER

Tan Tang · Junxiu Tang · Jiayi Hong · Lingyun Yu · Peiran Ren · Yingcai Wu

Design guidelines for augmenting short-form videos using animated data visualizations

Received: 8 July 2019 / Accepted: 4 August 2019 / Published online: 20 April 2020
© The Visualization Society of Japan 2020

Abstract Short-form videos are an increasingly prevalent medium for storytelling in journalism and marketing, of which information can be greatly enhanced by animated data visualizations. However, there is no prior research that systematically investigates how to augment such short videos with data visualizations in an effective way. We conducted a design workshop with experienced video, animation designers and visualization experts to discuss principles and practices for augmenting short-form videos with data visualizations. After the workshop, we summarized the participants' design considerations and proposed 20 design guidelines. We further collected design purposes of the participants and associated these purposes with the guidelines. Finally, we conducted a crowd-sourcing study and a task-based evaluation to validate the effectiveness and usability of the guidelines. Results indicate that our guidelines can significantly improve the videos accompanied with data visualizations and help novices easily obtain desired knowledge when augmenting videos.

Keywords Short-form video · Video augmentation · Design guidelines · Animated visualization · Data-driven storytelling

1 Introduction

Short-form videos are an efficient medium for communication and storytelling. However, existing short-form videos have heavily relied on textual information to express the underlying points, which lack the support of data visualizations. For example, commercial advertising videos introduce goods through subjective slogans, thereby leading to minimal persuasiveness. As "a picture is worth a thousand words," data visualization is an efficient method to enhance information of short-form videos and provide audiences additional objective insights. Furthermore, animated visualizations (Fisher 2010; Weng et al 2019) can also engage viewers, keep them oriented (Heer and Robertson 2007), and facilitate learning (Tversky et al 2002). Given these desirable features, short-form videos augmented with data visualizations, known as *data video* (Segel and Heer 2010), is a promising medium for next-generation presentation (Kosara and Mackinlay 2013).

T. Tang · J. Tang · J. Hong · Y. Wu (✉)
State Key Lab of CAD&CG, Zhejiang University and Alibaba-Zhejiang University Joint Institute of Frontier Technologies, Hangzhou, China
E-mail: ycwu@zju.edu.cn

L. Yu
Department of Computer Science and Software Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China

P. Ren
The Alibaba Group, Hangzhou, China

However, augmenting videos with data visualizations is complicated, especially when the video time is strictly constrained. Although data videos have become prominent in recent years (Amini et al 2015, 2017, 2018), no studies have systematically investigated the augmentation of data visualizations in short-form videos. The goals for the augmentation are clear: The videos should be augmented by data visualization; the data visualization should be able to capture the attention of viewers; the original information of the video should remain prominent. However, it is challenging to avoid the conflict between extra data visualization and original information in the video since both of them have to be noticed, understood and remembered in a short time. This core issue can be decomposed into several associated questions: **Q1** *What kind of data visualizations and animations can catch viewers' attention?* **Q2** *What positions of data visualizations in short-form videos can attract audiences?* **Q3** *How to craft impressive data visualizations of limited duration?* **Q4** *How to avoid conflict between data visualizations and original videos?* In particular, we aim to understand the design considerations of video editors when they address these questions.

To answer the above-mentioned questions, we conducted a design workshop and recruited 12 participants to perform a short-video augmentation task. We focused on adding data visualizations to existing videos, rather than crafting data videos from scratch. The task is crucial since there are many existing short-form videos, which need to be augmented by data visualization, especially in e-commerce websites. The participants were requested to consider the narrative of videos initially and then augment data visualizations in the videos. Our main goal is to propose design guidelines that enable visualizations to attract viewers without sacrificing the understanding of the original videos.

In the workshop, we observed that the participants relied on a common authoring process, that was, *what we want* and *how we do*. Thus, we conducted a post-interview to understand the design purposes of the participants. Although the design guidelines can capture the design knowledge of the participants (Van - Welie et al 2001; Xue et al 2019), these guidelines are too many and difficult to select. To improve the usability of the guidelines, we associated the guidelines with the design purposes and categorize the guidelines in accordance with the design purposes. Furthermore, we developed an interface to help novices in using these guidelines. Finally, we conducted a crowd-sourcing study to validate the effectiveness of the guidelines and a task-based evaluation to assess the usability of our interface.

The major contributions are summarized as follows:

1. We investigate a new problem of augmenting short-form videos with animated data visualizations and raise a set of research questions and propose a design space for follow-up studies.
2. We introduce a set of design guidelines and associated these guidelines with design purposes for video augmentation and design a tailored interface for users to adopt the guidelines easily.
3. We validate the effectiveness of the design guidelines through a crowd-sourcing study and prove the usability of the user interface with a laboratory study, the findings of which indicate that the guidelines can facilitate video augmentation for novice users.

2 Related work

2.1 Data-driven storytelling

Data-driven storytelling, also known as *narrative visualization* (Segel and Heer (2010)), has been regarded as the next step for visualization research (Kosara and Mackinlay 2013; Lee et al 2015; Xu et al 2019a). Segel and Heer first proposed a theoretical framework that describes seven genres of narrative visualizations (Segel and Heer 2010). Hullman and Diakopoulos (2011), Hullman et al (2013) further investigated the framing effect and narrative sequences of this emerging visualization. Tang et al (2019) extended the existing storyline algorithms to support more expressive narratives. Among the seven genres of narrative visualizations, data comics and data videos have become prominent in recent years (Amini et al 2015; Bach et al 2017). Bach et al (2017) introduced *data comics* and applied them to tell stories about dynamic networks (Bach et al 2016). Furthermore, they proposed design patterns to help storytellers craft data comics for general usage (Bach et al 2018). Amini et al (2015) first investigated data videos systematically and coded a set of videos using four narrative categories, namely *Establisher*, *Initial*, *Peak*, and *Release*. Moreover, they carefully examined the authoring process of data videos adapted by experienced storytellers. These researchers further developed a system that could help users create basic data motion graphics and construct units into a complete data story (Amini et al 2017). In contrast to their works, our studies focus on augmenting short-form videos with data visualizations. Specifically, we aim to explore the design guidelines that could help novice storytellers perform short-video augmentation effectively.

2.2 Animated data visualization

Animations are extensively used in data visualizations given its intuitive and engaging nature (Heer and Robertson 2007; Chevalier et al 2016; Zhou 2019). Researchers employed the animated transitions to demonstrate changes and differences in the consecutive states of dynamic graphs (Bach et al 2014; Rufiange and McGuffin 2013). For example, Bach et al (2014) proposed a visual interface with animated transitions to emphasize the changes in network structure and facilitate tracking and understanding of dynamic networks. Rufiange and McGuffin (2013) also developed a visual interface, called *DiffAni*, which adopts animation to demonstrate the evolution of graphs over time intervals. Furthermore, researchers investigated the considerable potential of animation in visualizing streaming data and have proposed various visual idioms (Huron et al 2013; Liu et al 2016; Wang et al 2016; Xu et al 2019b). For example, Huron et al (2013) presented a novel visual metaphor, called *Visual Sedimentation*, to visualize data streams collected from various disciplines. Visual sedimentation successfully adopts animation to solve the specific challenge of streaming visualizations, that is, to avoid visual clutter while keeping aging data visible. On the basis of this visual idiom, Liu et al (2016) proposed a real-time visual analytic method to help users explore and understand the topic evolution of sheer-volume text streams. Wang et al (2016) also adopted the *sedimentation* metaphor to develop a novel system that uses animated bubbles to represent video clickstream and presents accumulated click data with flow visualizations. Xie et al (2019) employed a smooth zooming animation to support multi-level exploration of image collections. Wu et al (2018) developed a dynamic line plot to visualize online social media data (Wu et al 2016). To improve audience engagement, we adopt the animated data visualizations to augment short-form videos.

2.3 Design guidelines in animations

Design guidelines refer to a set of suggestions toward a favorable design in practice (Foundation 2018). Many empirical studies have investigated the effectiveness of animated visualizations and have proposed various design guidelines (Heer and Robertson 2007; Amini et al 2018). Robertson et al (2008) found that using animations for trend visualizations is enjoyable and exciting. However, their findings indicated that animations may produce misleading information. Tversky et al (2002) suggested that two design principles, namely *Congruence* and *Apprehension*, must be followed to avoid such negative effects when crafting animated visualizations. According to the two high-level principles, Heer and Robertson (2007) proposed ten specific guidelines that focus on animated transitions in statistical data visualizations. Amini et al (2018) confirmed that pictographic representations and animations can enhance viewer engagement and developed four design guidelines for designing data videos. However, existing guidelines have insufficiently covered the entire design space of short-video augmentation. To compensate for the missing part, we conduct a workshop and invite 12 designers to share their design knowledge on short-video augmentation. On the basis of their responses, we propose modularized design guidelines to assist novices in augmenting short-form videos with data visualizations.

3 Video augmentation

Following the same idea of code augmentation (Hoffswell et al 2018), we aim to *place* data visualizations in short-form videos to enhance the understanding and present increasingly diverse information. Two *placement* methods are used for short-video augmentation. One is to layer data visualizations on the existing frames of videos (see Fig. 1c), and the other is to insert visualizations as a new frame in videos (see Fig. 1d). Our research has the following objectives: **O1** To understand the design considerations of short-video augmentation and propose design guidelines; **O2** To investigate design purposes and associate these purposes with the design guidelines. The first objective intends to investigate *how the participants do*, while the second objective aims to understand *what they want*.

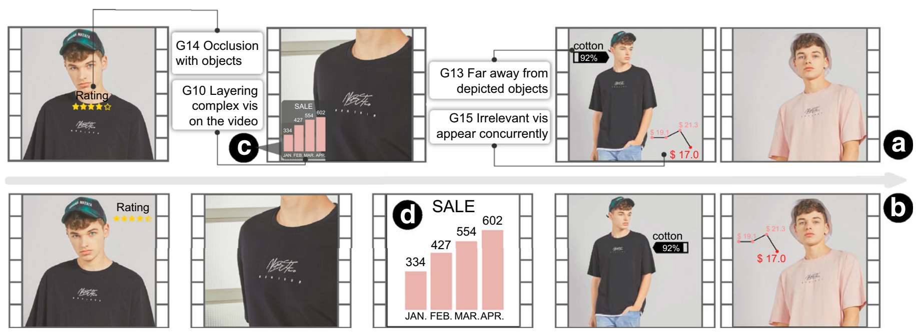

Fig. 1 **a** A short-form commercial video augmented with data visualizations which violate four design guidelines (G10, G13 to G15). **b** The video improved to follow these design guidelines. **c** Layering visualizations on an existing frame of the video. **d** Inserting visualizations as a new frame of the video

4 Design workshop

We conducted a design workshop to satisfy **O1** and **O2**. Similar studies have been successfully applied to identify cognitive design principles for assembly instructions (Heiser et al 2004; Li et al 2019) and understand foot gesture (Felberbaum and Lanir 2018; Xu et al 2017). We recruited designers who are experts in universal principles of design and have extensive experience in performing design tasks. These designers were first asked to augment short-form videos. Then, they were required to share their design considerations with affinity diagrams (Hanington and Martin 2012) (Fig. 2b–e). After that, we extracted the design considerations of the designers from these diagrams and summarized design guidelines of short-video augmentation (**O1**). Furthermore, we conducted a post-interview to understand the design purposes of the designers. We associated the guidelines with the design purposes and classify the guidelines to improve the retrieval efficiency (**O2**).

4.1 Participants

Twelve participants signed up for the workshop (4 female and 8 male, aged 21–26 years). They had formal training in design major including *Digital Media* (8), *Industrial Design* (2), and *Information Visualization* (4). Furthermore, these participants were skillful in editing videos and were familiar with basic tools, such as Adobe Premiere (Inc. 2003) and After Effects (Inc. 1993). All participants were familiar with data visualizations (mean = 3.25, range = 2-4 with 1 = "no experience" and 5 = "expertise"), and five of them had experiences in data videos. Each participant was paid 30$ for the 3-h workshop.

4.2 Study material

We focused on the design considerations of the participants when augmenting short-form videos, rather than pursuing high-quality data videos. We selected four short advertising videos from a popular e-commerce website[1]. The duration of videos ranges from 15 to 30 s (mean = 21.25 s). Furthermore, collecting the associated data about the products in the videos from the online website is also convenient. We obtained the price and attributes of the products, such as colors, materials, and size. We also provided customer ratings and short reviews about the products. The ratings consist of three measures, namely express, service, and quality, which are assessed using a five-point Likert scale. However, the participants were not limited by the provided data. They could use other data, such as the price of the competitive products, to perform the task.

4.3 Procedure

4.3.1 Phase 1: video augmentation

The workshop began with a 20-min introduction on the idea of short-video augmentation and data visualizations. Then, we presented two examples and four guiding questions (**Q1** to **Q4**) to illustrate their task

[1] https://www.taobao.com/

Fig. 2 Design considerations for augmenting short-form videos in the workshop. **a** Combined results after the inter-group discussion. **b–e** Inner-group discussion results

(10 min). Given the whole picture of the task, the participants were randomly divided into four groups. Furthermore, each group was assigned a video and its corresponding data. After that, they were asked to perform the task independently using any video editing tools on their own laptops. The task was limited to 60 min with emphasis on design ideas, rather than polished videos. Participants were allowed to submit their work with a delay of 24 h after the workshop.

4.3.2 Phase 2: group discussion

In this phase, we required all participants to share their design considerations on augmenting short-form videos with data visualizations. They were asked to first characterize the design space of video augmentations through an inner-group discussion and then map the results on a whiteboard (20 min). The whiteboard was divided into several regions to represent separate design dimensions of the design space. Afterward, the participants were required to write down their design considerations on stickers and put the stickers on the whiteboard (40 min). Finally, we organized an inter-group discussion to combine the results from the four groups (30 min). Similar design dimensions were merged, and stickers were reorganized in accordance with the new design space (Fig. 2a).

4.4 Post-interview

To further understand the design purposes of the participants, we conducted a post-interview. All participants were first asked to explain their stickers whose description and underlying ideas seem to be vague and unclear. Then, they were asked to explain *why* they come up with these stickers and *what* they want when augmenting short-form videos. We obtained six design purposes based on their responses.

4.5 Data processing

To satisfy **O1**, we collected 54 hand-written stickers from the workshop and transcribed them into electronic notes. Two researchers processed these notes independently. They first simplified the words of the notes to increase the clarity of the underlying idea. Second, they filtered notes that were unrelated to short-video augmentation. Third, they combined the notes that described similar ideas. Fourth, the researchers presented their results to each other and discussed to make an agreement on the results. Finally, we obtained 20 design guidelines on short-video augmentation with data visualizations from the notes (Fig. 3). To satisfy **O2**, we associated the guidelines with the design purposes that reveal the design requirements of the participants. For each design purpose, two researchers first related it to the guidelines separately and then discussed to reach an agreement on the result.

4.6 Results

In the workshop, we obtained 12 augmented videos that were used as the input to the next evaluation studies.

	Avoid Conflicts	Reduce Cognitive Load	Enhance Perception	Increase Appeal	Keep Consistency	Emphasize Data
Data						
G 1 Select the data that matches the intends of videos.	●					
G 2 Select appropriate visual idiom according to data characteristic.			●			
G 3 Use salient visual channels to encode important data attributes.			●			
G 4 Use semantic icons to improve the attraction of visualizations.				●		
G 5 Use colors which are in harmony with videos to embellish visualizations.				●	●	
Motion						
G 6 Apply simple and common animations to visualizations.			●			
G 7 Apply sequential animations to complicated visualizations.				●		
G 8 Apply the same animation to the composite glyphs of visualizations.			●			
G 9 Make animated visualizations quick and easy to follow.			●			
Layout						
G10 Insert complicated visualizations as a new frame in videos.		●				
G11 Separate visualizations from the informative frames of videos.		●				
G12 Determine the size of visualizations by the importance of data.			●			●
G13 Place visualizations next to the depicted objects in videos.			●	●		
G14 Avoid the occlusion between visualizations and the objects in videos.			●			
G15 Present irrelevant visualizations in different frames of videos.		●				
Duration						
G16 Determine the duration of visualizations according to the complexity.			●			
G17 Elongate the duration of important data visualizations.			●			●
G18 Shorten the duration of similar visualizations when displayed consecutively.			●			
Narrative						
G19 Present important data visualizations at the beginning or the end of videos.				●		●
G20 Arrange the order of visualizations to match the narrative structure of videos.	●				●	

Fig. 3 Design guidelines for video augmentation associated with design purposes

4.6.1 Design guidelines

We proposed 20 design guidelines that were organized by a design space of short-video augmentation. The design space consists of five dimensions, namely *data, motion, layout, duration*, and *narrative*, which are compliant with the guiding questions. In accordance with **Q1**, the first two dimensions reveal the way these participants presented data and animated visualizations. For **Q2**, the *layout* dimension indicates the considerations of the participants about the spatial positions of visualizations. For **Q3**, the *duration* dimension reveals the way to control the time of visualizations when presented in short-form videos. In relation to **Q4**, the *narrative* dimension suggests the methods to avoid the conflicts between the original narrative of videos and visualizations. All guidelines and their relationships with the design purposes are illustrated in Fig. 3. We only explain the guidelines that are innovative and abstract, because they are difficult to understand without further clarification.

The *data dimension* includes the selection and visualization of data, which is the start of the participants' tasks after watching the videos. This dimension comprises five guidelines, of which **G1** is explained below.

– *Select the data that matches the intent of videos* Data are an effective means of enhancing the information expressed in videos. To avoid confusing or misleading viewers, the data must serve the same goal of the videos. For example, the goal of advertising videos is to promote goods; thus, the selected data must exhibit a persuasive appeal to customers.

The *motion dimension* focuses on using an animation for presenting visualizations. Four guidelines contribute to this dimension, and we illustrate **G7** and **G8** here.

– *Apply sequential animations to complicated visualizations* Complicated visualizations refer to visual representations of complex data that contain numerous data items. To reduce the cognitive load of the viewers, an efficient method is to present data items sequentially (Heer and Robertson 2007).
– *Apply the same animation to composite glyphs of visualizations* Composite glyphs refer to the visual objects of the visualization which depict a collection of data items in a dataset (Borgo et al 2013). This guideline aims to avoid animating data visualizations improperly. For example, animating the rectangle glyphs of bar charts with different motions is messy for viewers. The guideline is compliant with the *readability* principle (Lidwell et al 2010a) which indicates that data visualizations should be easy to perceive and understand.

The *duration dimension* refers to the temporal distribution of animated visualizations. Three guidelines are with respect to this dimension, and we explain **G18** below.

- *Shorten the duration of similar visualizations when displayed consecutively* Priming effect (Lidwell et al 2010b) indicates that relevant concepts can be automatically activated in a working memory when similar data visualizations are presented consecutively. For short-form videos, limiting the duration of similar visualizations when they are displayed consecutively is efficient.

The *layout dimension* refers to the spatial positions of visualizations in videos. Six guidelines comprise this dimension, and four of them (**G10**, **G12**, **G13**, and **G15**) are illustrated.

- *Insert complicated visualizations as a new frame in videos* Complex visualizations may distract viewers' attention and hinder the understanding of the original videos. Thus, the complicated data must be visualized and displayed separately in videos.
- *Determine the size of visualizations by the importance of data* According to Fitts' law (Lidwell et al 2010c), an object is difficult to target when it is small and distant. Thus, the significance of the visualization must be decided by the importance of the data. This guideline aims to highlight data visualizations.
- *Place visualizations next to the depicted objects in videos* The *proximity* principle indicates that associated elements should be close to each other (Lidwell et al 2010d). Thus, visualizations should be placed near the objects which they described to improve the readability.
- *Present irrelevant data visualizations in different frames of videos* Interference effect (Lidwell et al 2010e) may occur when irrelevant datasets are visualized at the same time, thereby possibly confusing or even misleading viewers. Thus, separating irrelevant visualizations is important to avoid such information conflicts.

The *narrative dimension* refers to the sequence of visualization pieces, which is comprised of two guidelines.

- *Present important data at the beginning or the end of videos* Serial position effect (Lidwell et al 2010f) indicates that people can recognize the content at the beginning and end of a sequence. Thus, the guideline reveals an intelligent means of highlighting important data.
- *Arrange the order of visualizations to match the narrative structure of videos* Narrative structure refers to the order of story pieces in videos. In contrast to crafting a new video, video augmentations must first consider the narrative of the videos. Then, the order of data visualizations must be consistent with the narrative structure of the original videos.

4.6.2 Design purposes

We obtained six design purposes from the post-interview to understand the rationale of the guidelines.

- *Avoid conflicts* Visualizing data without carefully considering the original videos may produce conflicts, which refer to the mismatch between the insights of visualizations and the narrative (**G20**), or intent (**G1**) of videos. The participants aimed to avoid such conflicts for the visualizations to avoid disturbing the original information of videos.
- *Enhance perception* Perception is the process of organizing, identifying, and interpreting conceived visual information. The participants aimed to enhance the perception of data visualizations (**G2** and **G3**) and animations (**G6**, **G8**, and **G9**) for easily noticing and understanding. This design purpose is also facilitated by the guidelines from the *duration* (**G16** to **G18**) and *layout* (**G12** to **G14**) dimensions.
- *Increase appeal* Visual appeal is important in creating a good first impression and evokes positive emotion for audiences. The participants aimed to increase the appeal of visualizations (**G4** and **G5**) to attract viewers' attention and inspire their interest in videos. Findings from the study (Heer and Robertson 2007) indicate that sequential animations (**G7**) could increase the appeal of visualizations, which is also facilitated by the guidelines from *layout* (**G13**) and *narrative* (**G19**).
- *Reduce cognitive load* Cognitive load is determined by the complexity of the perceived information. The participants aimed to reduce the cognitive load for the visualizations to be easily interpreted. This design purpose is in regard to the usage of the two visualization *placement* methods which are related to the *layout* dimension (**G10**, **G11**, and **G15**).

– *Emphasize data.* Data are efficient ways of enhancing the textual information in videos. The participants used visualizations to highlight data that can support the main idea of the videos. This design purpose could be facilitated by increasing the size (**G12**), elongating the duration (**G17**), and considering the order (**G19**) of visualizations.
– *Keep consistency* The *consistency* principle (Lidwell et al 2010g) indicates that the similar components in a system should have consistent appearance or behaviors. The participants extended this universal principle in the context of video augmentations. The visualizations to be augmented in videos must be consistent with the style (**G5**) and content (**G20**) of the original videos to avoid producing abrupt impressions.

5 User interface

In this section, we introduce a user interface to enable novice users to fully utilize the proposed guidelines. To make the guidelines easy to access and navigate, we employ a list view to provide a quick overview (Fig. 4a). For each guideline, we develop a collapsed card (Fig. 4d) which presents thumbnail information for users to support the level-of-detail exploration. Users can expand the collapsed card to obtain comprehensive information about the selected guideline. To make the guidelines easy to select, we provide a navigation bar (Fig. 4b) which presents the dimensions of the design space and purposes using two dropmenus (Fig. 4c). Users can click an option in the dropmenus to select a specific design dimension or design purpose, and the list view will only present the associated guidelines.

6 Evaluation

In this section, we conducted a crowd-sourcing study to validate the effectiveness of the design guidelines. Crowd-sourcing studies have been successfully practiced in assessing graphic designs (Heer and Bostock 2010) and generating visualizations (Singh et al 2018). We also conducted a task-based evaluation to assess the usability of the guidelines.

6.1 Crowd-sourcing study

6.1.1 Study design

We obtained 12 short-form videos accompanied with data visualizations in the design workshop. These videos were not polished due to the lack of the design guidelines. We selected these visual recordings as baseline videos. In order to investigate whether these videos can be improved by the short-video

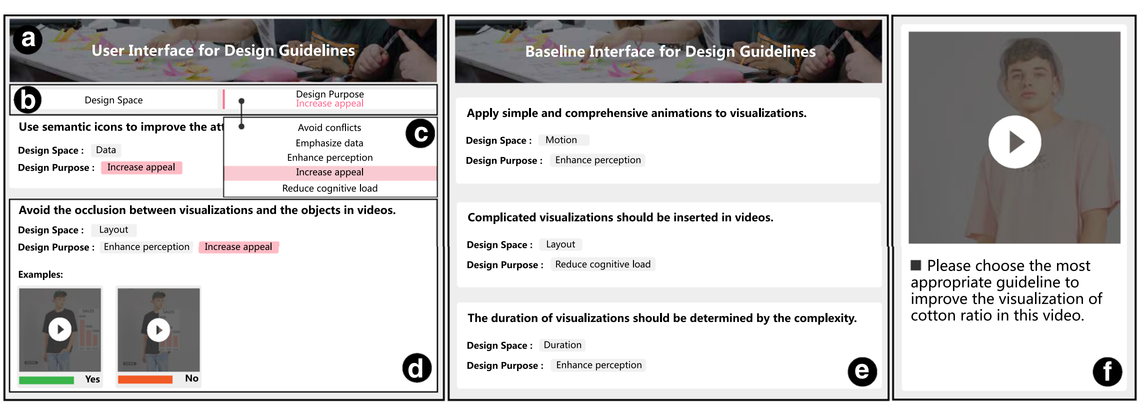

Fig. 4 **a** Our interface demonstrates guidelines with a list view. **b** A navigation bar consists of two dropmenus **c** which can help novices easily select guidelines. **d** A collapsible card can present thumbnail or comprehensive information of each guideline. **e** The baseline interface of the usability study can only present the whole list of guidelines. **f** A video card presents the task information

augmentation guidelines, we conducted a comparison crowd-sourcing study. We first invited two external experts to evaluate the baseline videos according to the design guidelines and then improved these videos based on the experts' comments. Crowdworkers were required to complete tasks and provide subjective ratings to both the baseline and improved videos. The effectiveness of the design guidelines was validated by comparing the task accuracy and subjective ratings of the two video groups. A similar idea was employed in an empirical study of the web-form improvement guidelines (Seckler et al 2014). The study was deployed on *Amazon Mechanical Turk* (Amazon 2005).

6.1.2 Stimuli

The stimuli consist of two video groups, namely baseline videos and improved videos. The baseline videos were crafted by the participants in the design workshop and were later improved by two researchers through the following phases.

Phase 1 We invited two external experts with more than ten years of experiences in UI/UX design to evaluate the 12 short-form videos obtained in the workshop. Both experts were collaborators of this project and familiar with the idea of short-video augmentation and the design guidelines. They were asked to evaluate the 12 videos according to the design guidelines. For each guideline, the experts were asked to score the videos using a 5-point Likert scale ("1" = mismatch the guideline and "5" = match the guideline very well). For the guideline with score of less than 3, the experts were required to comment on how to improve the corresponding videos. The experts first completed evaluations independently, and then discussed their results for the final decisions.

Phase 2 Four video categories were considered in the design workshop (see Fig. 5). We defined the overall score of videos as the mean value of all gradings and selected four baseline videos with the lowest grading in each category (see Fig. 5). Two researchers then proposed and discussed improvement proposals for the selected videos according to the experts' comments. After reaching an agreement on the final proposals, the researchers improved the videos and invited the experts to re-evaluate the results. Improvements were applied iteratively until the experts were satisfied. Lastly, we obtained the four improved videos. In total, we had eight short-form videos (four baseline and four improved videos) as stimuli for the study.

6.1.3 Questionnaire

At the beginning of the questionnaire, we introduced the basic ideas of data visualization. To evaluate the performance of the stimuli, we proposed six questions for each video, including two task questions (Fig. 6b) and four subjective questions (Fig. 6a). The two task questions provided objective measurements. Task 1

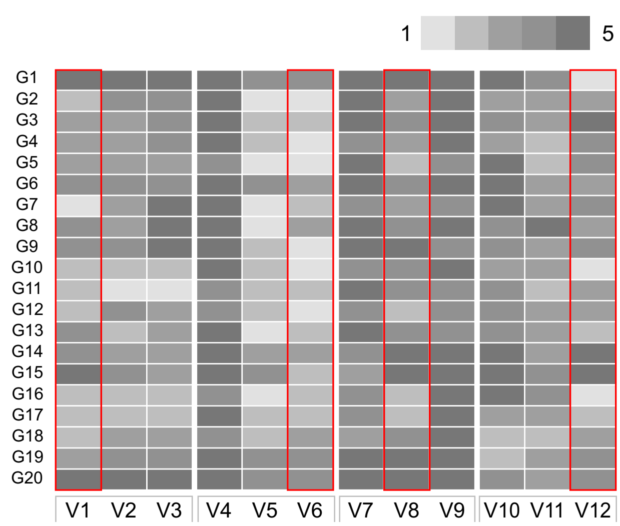

Fig. 5 The review results of the short-form videos obtained in the workshop by external experts. The rows represent guidelines G1 to G20, while the columns represent videos V1 to V12 which belong to four categories. The gray value of cells present scores 1 to 5. Red rectangles indicate videos are selected for the crowd-sourcing study

aimed to determine whether data visualizations hinder the delivery of original information, whereas Task 2 aimed to understand whether data visualizations can capture the attention of viewers. For example, crowdworkers were asked to identify the brand (Task 1) and sales (Task 2) of the products shown in screened videos. We randomly assigned one Task 1 question and one Task 2 question to the videos. The four subjective questions aimed to understand the effects of data visualizations from two aspects: legibility and aesthetics. These subjective questions were answered using a 7-point Likert scale ("1" = strongly disagree; "7" = strongly agree). The videos appeared in a random order. In total, crowdworkers were required to complete 48 questions for each questionnaire.

6.1.4 Results

In the study, we assigned 130 HITs to crowdworkers and received 86 responses. We filtered the questionnaires with the completion time of less than 5 min (videos time 2.6 min and default question time 2.4 min) and obtained 85 valid responses. To understand the effects of improvement, we obtained the quantitative results on both tasks (Fig. 6b) and subjective ratings (Fig. 6a). For the total task questions and ratings, we used paired t-test with 95% confidence interval and observed the significantly better performance of the improved videos in terms of task accuracy ($p = 0.006562$) and ratings ($p < 0.001$) compared with the baseline videos.

For the task accuracy, the improved videos performed better in Tasks 1 and 2 than the baseline videos. In Task 1, the crowdworkers were required to recall the objects shown in the stimuli. Figure 6b indicates that both improved and baseline videos exhibit good performance ($p = 0.052$, accuracy $= 91.3\%$ and 84.7%, respectively) in this task. The possible reason is that these videos avoid hindering original video information given their good scores for the guidelines of the *narrative* dimension (see Fig. 5). In particular, guideline **G20** aims to avoid conflicts between visualizations and videos, facilitating the understanding of the original information. We observed that the improved videos (accuracy $= 74.8\%$) exhibited a significantly better performance ($p = 0.022$) than the baseline videos (accuracy $= 52.7\%$) in Task 2, which required the crowdworkers to identify data patterns in visualizations. We infer that possibly the baseline videos violated guidelines **G16** and **G17**. The duration of visualizations in these videos is too short, causing difficulty to catch data information delivered by visualizations.

Figure 6a depicts the subjective ratings of how easy data visualizations were to notice and understand, how well they fit the videos, and whether they were visually appealing. We found a significantly positive effect for the improved videos in easy to notice ($p < 0.001$) and understand ($p < 0.001$), which indicated that it was easier for the crowdworkers to perceive visualizations in the improved videos. We found the improved videos also had a significantly better performance in visual appealing of data visualizations ($p < 0.001$). The data visualizations in the improved videos also perform significantly better in the fitness of the original videos than those in baseline condition ($p < 0.001$).

6.2 Usability study

6.2.1 Study design

We conducted a task-based evaluation to validate the usability of short-video augmentation guidelines. Specifically, we aimed to investigate whether our interface can enable novice users to easily select proper

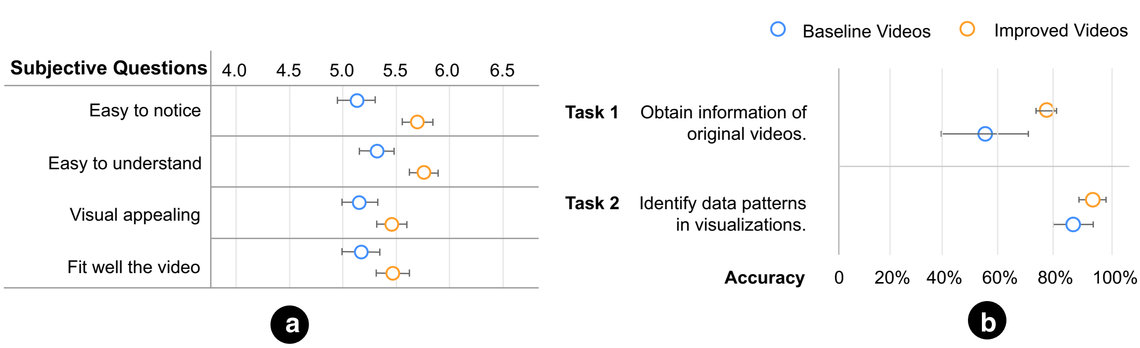

Fig. 6 The **a** subjective ratings and **b** task accuracy of the crowd-sourcing study with 95% confidence interval

guidelines when augmenting short-form videos. We proposed 15 tasks, and participants used either our interface or baseline interface to select guidelines which solve the issues of short-video augmentation introduced in the tasks. Our assumption is that users can easily locate correct guidelines by reducing the search space through selecting design dimensions or purposes. Usability was measured by task completion time and accuracy. Short interviews were conducted to further understand user satisfaction and improvement suggestions.

6.2.2 Participants

To understand the usage of guidelines for novice users, we recruited 24 participants (12 male and 12 female) from various backgrounds, including computer science, math, social science, and biological engineering. Their average age was 23 years old (range = 19–26 years). Twelve participants used the baseline interface to complete tasks, whereas the other participants performed the tasks using our interface. Independent Student's t-tests indicated no significant differences between the two groups with regard to age, computer knowledge, and visualization expertise. Each participant was paid 10$ for the 30-min study.

6.2.3 Baseline interface

In order to compare with our user interface, we developed a baseline interface (Fig. 4e). The baseline interface removes the dropmenus of the design space and purposes, ensuring that users can only enumerate the guidelines to find the desired one. To better demonstrate the tasks, we enhanced the two interfaces with a video card (see Fig. 4f), which consists of a task video and instructions.

6.2.4 Tasks

We proposed 15 tasks to validate the usability of the design guidelines. Each task contained a 5-s video which was extracted from the source videos in the design workshop. We added extra data visualization to the task videos, and each video was designed to deliberately violate one short-video augmentation guideline. In the task, the participants were asked to identify the violated guideline to help them improve the screened video. The tasks were divided into two groups: 3 training tests and 12 formal tests.

6.2.5 Procedure

In the training phase (10 min), we first provided the participants with general information about the study. Then, the participants were informed to complete three training tests. We accompanied the participants to answer their questions until they became familiar with the tasks. After that, the participants started the formal tests (20 min), with a timer recording the task completion time. The tasks were assigned to the participants in a random order. Half of the participants can select guidelines from the list in the baseline mode, and the others were allowed to use the navigation bar from the interface in experimental mode. After completing the formal tasks, the participants were interviewed on "whether the interface is easy to use or select guidelines" and "what suggestions they have to further improve the usability."

6.2.6 Results

We obtained the accuracy and completion time of each task. Figure 7 indicates that the average accuracy of the 12 formal tests completed by our interface (accuracy = 95.2%) is higher than the baseline condition

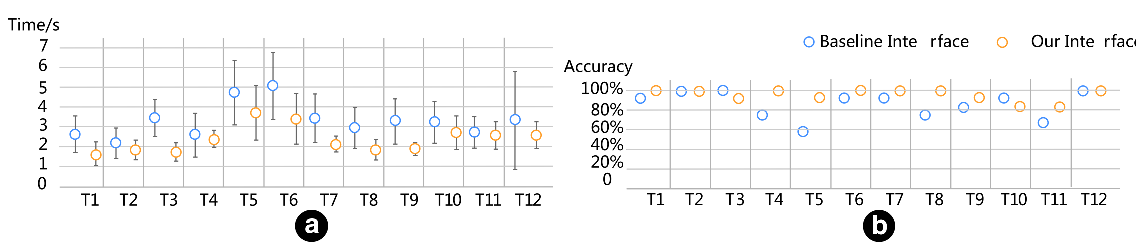

Fig. 7 The **a** completion time and **b** task accuracy of the usability study with 95% confidence interval

(accuracy = 85.5%). We used a paired t-test and identified a significantly positive effect for our interface in terms of the task accuracy ($p = 0.0145$). We also observed the significantly negative effect of our interface on the completion time of tasks ($p < 0.001$). The quantitative results indicate that classifying short-video augmentation guidelines according to the design space and purposes can significantly improve the performance with regard to the task accuracy and completion time. Figure 7 also shows two special cases (T3 and T10) where the baseline interface shows a better performance in terms of the task accuracy than our interface. This result is attributed to the considerable time spent by participants when completing these tasks to increase the chance of obtaining right answers.

7 Discussion

In this section, we discuss the implications and limitations of our studies, along with the generability and usability of video augmentation guidelines.

Design implications This research is the first time to understand how people perceive extra data visualization from a short video. The limited time of short-form videos causes difficulty in perceiving expressive information in visualizations among audiences. Thus, the existing content in short-form videos must be considered when adding extra data visualizations. However, no previous work has systematically investigated how visualizations interact with the original video content. Thus, we conducted a design workshop to answer four important research questions (**Q1** to **Q4**) and proposed a set of short-video augmentation guidelines. Findings from the crowd-sourcing study indicate that our guidelines can successfully improve the short commercial advertising videos. Crowdworkers can easily identify data patterns in visualizations while noticing the original information of videos.

We further discuss the design implications of other findings in the workshop. We distill a design space of short-video augmentation; this space consists of five design dimensions(*data, motion, layout, duration*, and *narrative*). The design space inspires a nonlinear iterative authoring process which can guide novice users to augment videos with our guidelines efficiently. *Step 1* start video augmentation tasks by selecting and visualizing *data* after familiarizing with videos; *Step 2* use proper *motion* to present visualizations and decide the *duration*; *Step 3* consider the spatial *layout* of the videos and select appropriate positions to place visualizations; *Step 4* contemplate whether the visualizations can match the original *narrative* of videos, otherwise retrace the last three steps to re-design the visualizations. This design process can also inspire an authoring tool which facilitates short-video augmentation with an efficient workflow.

Generability Through the workshop, we propose a set of design guidelines to facilitate short-video augmentation. However, the workshop is established on the basis of commercial advertising videos that primarily aim to promote products and convince audiences. For other video types with different intents, such as news and public service videos, whether our guidelines are sufficient or applicable remains unknown. Moreover, applying the guidelines to videos with no time limitations also requires further studies. As a next step, we plan to extend our research on a broader range of videos. Specifically, we plan to invite more participants from various industries including journalism, public-interest groups, and government agencies to extend our guidelines.

Usability Two obstacles prevent the full utilization of the design guidelines for video augmentation. The participants from the baseline group in the usability study mentioned: (1) The guidelines are too many to select, and (2) several guidelines are difficult to interpret without further clarification. To improve the interpretability, researchers (Van Welie et al 2001; Bach et al 2018) proposed design patterns to code guidelines from various aspects including design problems, principles, and context. However, the increasing complexity of guidelines leads to higher learning costs and lower retrieval efficiency. The coding method should be carefully designed to achieve the trade-off between interpretability and simplicity. Furthermore, the guidelines should be classified to reduce the search space and improve the retrieval efficiency. We associate the guidelines with the design purposes collected from the post-interview in the workshop. The design purposes indicate the goals of the participants when augmenting short videos, revealing the design requirements for video editors. We further classify the guidelines according to the design purposes to enable users to locate the guidelines promptly according to their demands. Results of the usability study indicate that our interface performed better in both task accuracy and completion times.

The participants from the experimental group also provide valuable suggestions. They mentioned that the design purposes were highly advanced for novice users. “It is better to support users search guidelines according to a specific design problem.” We leave this issue for future work. We plan to survey specific design problems of video augmentation and associate them with design purposes to allow novices to directly search guidelines through design problems.

Study limitations As a preliminary evaluation, the crowd-sourcing study aimed to assess the effectiveness of short-video augmentation guidelines. We obtained four baseline videos crafted by the participants in the workshop and improved them to satisfy all guidelines. However, the crowd-sourcing study only validated the combined effectiveness of the guidelines on commercial advertising videos. Moreover, only the guidelines violated by the baseline videos (see Fig. 5) were used to guide the improvement. As a next step, we plan to further evaluate the single effectiveness of each guideline on a broader range of videos. In the usability study, we recruited participants who were novices in video editing and augmentation from an internal database of our institute. Although the study results indicate that our interface can significantly improve the use of guidelines, we acknowledge the potential bias considering the users' background. Thus, we plan to extend the study to a broader range of participants on different ages and educations levels.

8 Conclusion

In this paper, we investigate how to enhance information by using data visualization in the short-form videos. We characterize four associated research questions (**Q1** to **Q4**) about this issue and conduct a design workshop to answer these questions. Through the workshop, we distill a design space of short-video augmentation and propose 20 design guidelines to help novices augment short-form videos. We further collect design purposes of the guidelines from the participants in the workshop to enable novices to easily select guidelines according to their demands. Finally, we conduct a crowd-sourcing study and an in-laboratory study to validate the effectiveness and usability of the guidelines.

Several avenues exist for future work. As the first step to understand how to augment short-form videos with data visualizations, we plan to extend our guidelines on a broader range of videos. We also plan to associate the guidelines with specific design problems to further improve the usability for novices. Moreover, the design guidelines only facilitate human users to augment short-form videos. As a next step, we intend to investigate how to employ these guidelines in machine learning models to develop automatic tools that augment short-form videos efficiently.

Acknowledgements The work was supported by NSFC (61761136020), NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (U1609217), Zhejiang Provincial Natural Science Foundation (LR18F020001) and the 100 Talents Program of Zhejiang University. This Project was also supported by Alibaba Group.

References

Amazon (2005) Amazon mechanical turk. https://www.mturk.com/. Accessed 1 Aug 2018

Amini F, Henry Riche N, Lee B, Hurter C, Irani P (2015) Understanding data videos: Looking at narrative visualization through the cinematography lens. In: Proceedings of the ACM conference on human factors in computing systems, pp 1459–1468

Amini F, Riche NH, Lee B, Monroy-Hernandez A, Irani P (2017) Authoring data-driven videos with dataclips. IEEE Trans Vis Comput Graph 23(1):501–510

Amini F, Riche NH, Lee B, Leboe-McGowan J, Irani P (2018) Hooked on data videos: assessing the effect of animation and pictographs on viewer engagement. In: Proceedings of the international conference on advanced visual interfaces, pp 1–9

Bach B, Pietriga E, Fekete J, Bach B, Pietriga E, Fekete J, Member S (2014) Navigation for dynamic networks graphdiaries: animated transitions and temporal navigation for dynamic networks. IEEE Trans Vis Comput Graph 20(5):740–754

Bach B, Kerracher N, Hall KW, Carpendale S, Kennedy J, Henry Riche N (2016) Telling stories about dynamic networks with graph comics. In: Proceedings of the ACM conference on human factors in computing systems, pp 3670–3682

Bach B, Riche NH, Carpendale S, Pfister H (2017) The emerging genre of data comics. IEEE Comput Graph Appl 38(3):6–13

Bach B, Wang Z, Farinella M, Murray-Rust D, Henry Riche N (2018) Design patterns for data comics. In: Proceedings of the ACM conference on human factors in computing systems, pp 1–12

Borgo R, Kehrer J, Chung DHS, Maguire E, Laramee RS, Hauser H, Ward M, Chen M (2013) Glyph-based visualization: foundations, design guidelines, techniques and applications. Eurographics State of the Art Reports, pp 39–63

Chevalier F, Riche N, Plaisant C (2016) Animations 25 Years Later: New Roles and Opportunities. In: Proceedings of the international working conference on advanced visual interfaces, pp 280–287

Felberbaum Y, Lanir J (2018) Better understanding of foot gestures: an elicitation study. ACM, New York, p 334

Fisher D (2010) Animation for visualization: opportunities and drawbacks. In: Beautiful visualization. O'Reilly Media, pp 329–352

Foundation ID (2018) Design guidelines. https://www.interaction-design.org/literature/topics/design-guidelines. Accessed 1 Aug 2018

Hanington B, Martin B (2012) Affinity diagramming. In: Universal methods of design, pp 12–14. Rockport Publishers

Heer J, Bostock M (2010) Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In: Proceedings of the ACM conference on human factors in computing systems, ACM, pp 203–212
Heer J, Robertson GG (2007) Animated transitions in statistical data graphics. IEEE Trans Vis Comput Graph 13(6):1240–1247
Heiser J, Phan D, Agrawala M, Tversky B, Hanrahan P (2004) Identification and validation of cognitive design principles for automated generation of assembly instructions. In: Proceedings of the working conference on advanced visual interfaces, ACM, pp 311–319
Hoffswell J, Satyanarayan A, Heer J (2018) Augmenting code with in situ visualizations to aid program understanding. In: Proceedings of the 2018 CHI conference on human factors in computing systems, ACM, p 532
Hullman J, Diakopoulos N (2011) Visualization rhetoric: framing effects in narrative visualization. IEEE Trans Vis Comput Graph 17(12):2231–2240
Hullman J, Drucker S, Henry Riche N, Lee B, Fisher D, Adar E (2013) A deeper understanding of sequence in narrative visualization. IEEE Trans Vis Comput Graph 19(12):2406–2415
Huron S, Vuillemot R, Fekete JD (2013) Visual sedimentation. IEEE Trans Vis Comput Graph 19(12):2446–2455
Inc A (1993) Adobe after effects. https://www.adobe.com/products/aftereffects.html. Accessed 1 Aug 2018
Inc A (2003) Adobe premiere. https://www.adobe.com/products/premiere.html. Accessed 1 Aug 2018
Kosara R, Mackinlay J (2013) Storytelling: the next step for visualization. IEEE Comput 46(5):44–50
Lee B, Riche NH, Isenberg P, Carpendale S (2015) More than telling a story: transforming data into visually shared stories. IEEE Comput Graph Appl 35(5):84–90
Li C, Lv P, Manocha D, Wang H, Li Y, Zhou B, Xu M (2019) ACSEE: antagonistic crowd simulation model with emotional contagion and evolutionary game. In: IEEE transactions on affective computing (**to appear**)
Lidwell W, Holden K, Butler J (2010a) Readability. In: Universal principles of design, pp 198–199. Rockport Publishers
Lidwell W, Holden K, Butler J (2010b) Priming. In: Universal principles of design, pp 186–187. Rockport Publishers
Lidwell W, Holden K, Butler J (2010c) Fitts' Law. In: Universal principles of design, pp 98–99. Rockport Publishers
Lidwell W, Holden K, Butler J (2010d) Proximity. In: Universal principles of design, pp 196–197. Rockport Publishers
Lidwell W, Holden K, Butler J (2010e) Inference effects. In: Universal principles of design, pp 138–139. Rockport Publishers
Lidwell W, Holden K, Butler J (2010f) Serial position effect. In: Universal principles of design, pp 220–221. Rockport Publishers
Lidwell W, Holden K, Butler J (2010g) Consistency. In: Universal principles of design, pp 56–57. Rockport Publishers
Liu S, Yin J, Wang X, Cui W, Cao K, Pei J (2016) Online visual analytics of text streams. IEEE Trans Vis Comput Graph 22(11):2451–2466
Robertson G, Fernandez R, Fisher D, Lee B, Stasko J (2008) Effectiveness of animation in trend visualization. IEEE Trans Vis Comput Graph 14(6):1325–1332
Rufiange S, McGuffin MJ (2013) DiffAni: visualizing dynamic graphs with a hybrid of difference maps and animation. IEEE Trans Vis Comput Graph 19(12):2556–2565
Seckler M, Heinz S, Bargas-Avila JA, Opwis K, Tuch AN (2014) Designing usable web forms—empirical evaluation of web form improvement guidelines. In: Proceedings of the ACM conference on human factors in computing systems, pp 1275–1284
Segel E, Heer J (2010) Narrative visualization: telling stories with data. IEEE Trans Vis Comput Graph 16(6):1139–1148
Singh DP, Lisle L, Murali T, Luther K (2018) Crowdlayout: crowdsourced design and evaluation of biological network visualizations. In: Proceedings of the ACM conference on human factors in computing systems, ACM, pp 1–14
Tang T, Rubab S, Lai J, Cui W, Yu L, Wu Y (2019) iStoryline: effective convergence to hand-drawn storylines. IEEE Trans Vis Comput Graph 25(1):769–778
Tversky B, Morrison JB, Betrancourt M (2002) Animation: Can it facilitate? Int J Hum Comput Stud Schnotz Kulhavy 57:247–262
Van Welie M, Van Der Veer GC, Eliëns A (2001) Patterns as tools for user interface design. In: Tools for Working with Guidelines, pp 313–324
Wang Y, Chen Z, Ma X, Luo Q, Qu H (2016) Animated Narrative Visualization for Video Clickstream Data. In: Proceedings of Siggraph Asia symposium on visualization, pp 1–8
Weng D, Chen R, Deng Z, Wu F, Chen J, Wu Y (2019) SRVis: towards better spatial integration in ranking visualization. IEEE Trans Vis Comput Graph 25(1):459–469
Wu Y, Cao N, Gotz D, Tan YP, Keim DA (2016) A survey on visual analytics of social media data. IEEE Trans Multimed 18(11):2135–2148
Wu Y, Chen Z, Sun G, Xie X, Cao N, Liu S, Cui W (2018) StreamExplorer: a multi-stage system for visually exploring events in social streams. IEEE Trans Vis Comput Graph 24(10):2758–2772
Xie X, Cai X, Zhou J, Cao N, Wu Y (2019) A semantic-based method for visualizing large image collections. IEEE Trans Vis Comput Graph 25(7):2362–2377
Xu M, Li C, Lv P, Lin N, Hou R, Zhou B (2017) An efficient method of crowd aggregation computation in public areas. IEEE Trans Circuits Syst Video Technol 28(10):2814–2825
Xu M, Su H, Li Y, Li X, Liao J, Niu J, Lv P, Zhou B (2019a) Stylized aesthetic QR code. IEEE Trans Multimed 21(8):1960–1970
Xu M, Xie X, Lv P, Niu J, Wang H, Li C, Zhu R, Deng Z, Zhou B (2019b) Crowd behavior simulation with emotional contagion in unexpected multihazard situations. In: IEEE transactions on systems, man, and cybernetics: systems (**to appear**)
Xue J, Yin H, Lv P, Xu M, Li Y (2019) Crowd queuing simulation with an improved emotional contagion model. Sci China Inf Sci 62(4):44101
Zhou Z (2019) Abductive learning: towards bridging machine learning and logical reasoning. Sci China Inf Sci 62(7):76101

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.