VisJockey: Enriching Data Stories through Orchestrated Interactive Visualization*

Bum Chul Kwon
University of Konstanz
Konstanz, Gemrany
bumchul.kwon@uni-konstanz.de

Florian Stoffel
University of Konstanz
Konstanz, Gemrany
florian.stoffel@uni-konstanz.de

Dominik Jäckle
University of Konstanz
Konstanz, Gemrany
dominik.jaeckle@uni-konstanz.de

Bongshin Lee
Microsoft Research
Redmond, WA, USA
bongshin@microsoft.com

Daniel Keim
University of Konstanz
Konstanz, Gemrany
daniel.keim@uni-konstanz.de

ABSTRACT

Visualization has recently begun to extend its role into a communication medium for presentation and storytelling. With the advancement in web-based visualization technology (e.g., D3), we increasingly encounter an integration of interactive visualizations into data stories in news media, blog posts, etc. However, these stories usually do not provide enough guidance on how to interpret and manipulate the accompanied visualizations. Therefore, readers are often on their own in finding the right state and area of visualization authors intended to show to support their arguments. We introduce *VisJockey*, a technique that enables readers to easily access authors' intended view through orchestrated visualization. To offload readers' burden in making connections between the text and the visualization, VisJockey augments the visualization through highlight, annotation, and animation. We describe the main concept of VisJockey and present three example stories.

Categories and Subject Descriptors

H.5.0 [**Information Interfaces and Presentations**]: General

General Terms

Design

Keywords

Visualization; Storytelling; Data story; Text link; Journalism; Web

1. INTRODUCTION

Many online news organizations (e.g., New York Times, Guardian, and Washington Post) have been increasingly integrating visualizations into their articles. With a tight script, these visualizations can help to communicate complex ideas and provide clear explanations. Furthermore, they give readers an opportunity to explore the data and to discover their own findings that may not be mentioned in the article. With the recent advancement in web-based visualization technology (e.g., D3), these articles often introduce new, custom visualizations to best match their topic.

However, such visualizations place extra burden on readers because they are usually introduced without proper instruction. If readers have not seen the visualization before, they need to figure out how to interpret it by themselves. Furthermore, they are usually on their own in making connections between the text and visualization. Since this can be challenging even for the commonly used, static bar charts [7], it would be much more difficult for readers to find the right state (or configuration) and area of visualization authors intended to show to support their arguments. Therefore, visualizations could be substantially underused or even misused, hindering readers from understanding the article.

To address this issue, we explore ways to enable readers to fully leverage the accompanied visualization while offloading their burden by tightly coupling the text with embedded visualizations. In this paper, we present VisJockey (Figure 1), a technique developed to help readers access authors' intended view more easily by illustrating the relevant text segments with visualizations. VisJockey invokes the interactions the visualization already supports (e.g., select and filter) and provides its own visual effects–annotation, highlight, and animated transition. The combination of interactions and visual effects is an action that can be preloaded and then played later on demand to help readers make the connection between the text and visualization. To indicate that an illustration with visualization is available without interrupting reading, VisJockey places small symbolic icons inline (e.g., ▦) and highlights the text segment on mouse over (Figure 1). Readers could invoke this guided action by interacting with the text segment. For example, when a reader clicks on the linked text identified with an icon, the visualization plays the preloaded action, which selects four countries and performs the animation, matching the text description.

We organize this paper as follows. We first describe the recent work related to the use of visualization as a storytelling medium, and explain the main concept of VisJockey. To demonstrate the applicability of the VisJocky technique, we provide three examplary
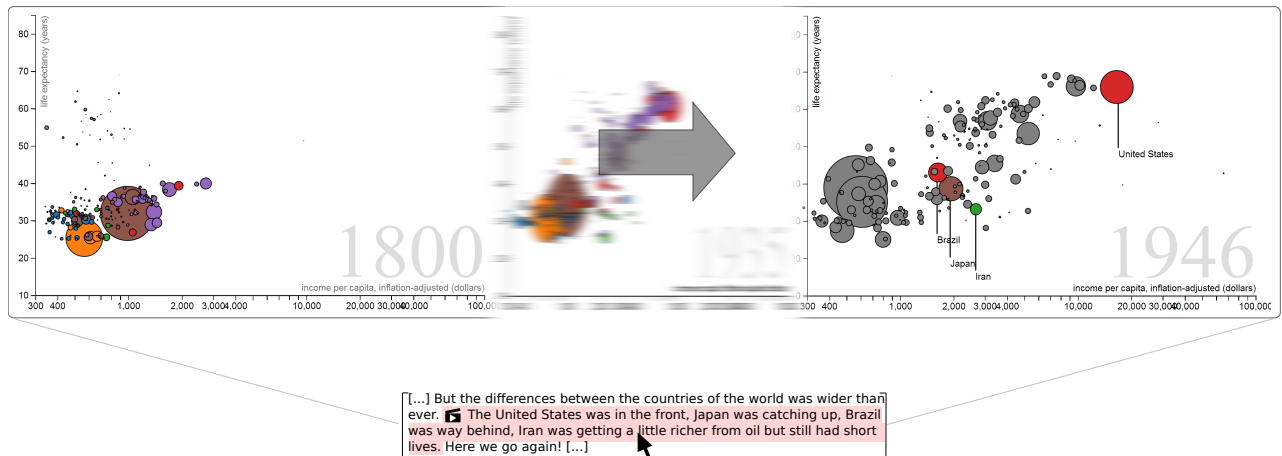

Figure 1: An illustration of a play of a preloaded action with VisJockey. When readers click on linked text segment, VisJockey shows animated transitions of the bubble chart.

stories using our technique. Finally, we conclude with discussion and future direction for improvements.

2. RELATED WORK

Even though storytelling has only recently become a focus of research in the visualization community, many visualization systems have long been used to support and enhance communication between people. One classic example is Many Eyes [15], which provides a website where people can easily construct visualizations with their own data and share findings through various social features. CommentSpace facilitates embedded chart-based discussions by enabling users to tag and link user-generated comments to organize findings [16]. These tools can be useful for many people to collectively analyze data and discuss insights using visualizations. Recently, interactive visualizations have increasingly been integrated into online articles (e.g., [12]).

However, we often encounter a situation, where readers cannot understand authors' intention properly because it is difficult to "reproduce interactive features in a static medium" [4, p. 99]. Segel and Heer [14] note that current news visualizations seem to lack commentary and narratives that can highlight key findings in visualizations. Thus, storytelling using visualizations may require more than visualization skills to integrate various components of stories like text narrative [3]. Since complex visualizations may overwhelm the general public with misinterpretation and confusion, narratives and annotations can play an important role in lowering the entrance bar [11]. Therefore, text narratives may complement to interactive visualizations to increase the readability.

Several recent studies investigate methods to connect text and visualizations. For example, Contextifier automatically generates annotations on visualizations used for news articles [6]. Kong et al. investigated a method to find connections between text and line charts in news articles using crowdsourcing [7]. line is a technique that places a small visualization within inline text so that people can understand the gist of it with simple patterns [2]. VisJockey can be combined with these methods to better support asynchronous communication, allowing authors to fully leverage visualizations to support their arguments.

While less common than the asynchronous storytelling, there are several tools to support live storytelling where both sides of communication present at the same time. For example, SketchStory is specifically designed to enable presenters to engage the audience by letting them interactively draw charts during the live presentation [10]. Tableau recently introduced a feature called Story Points, which aims to capture interesting visualization states and integrate them with a storyline annotated by presenters, which can be shown during the live presentation [1]. VisJockey aims to fill in for absent presenters in asynchronous communication by encoding and illustrating their intended actions.

In addition, high-level discussion of narrative visualization has started in the community. Segel and Heer [14] investigate how narrative visualization differs from the traditional form of storytelling. They review the design space with infographics in the context of data journalism and identify three divisions of features-genre, visual narrative, and narrative structure. Hullman and Diakopoulos [5] explore how the rhetorical conventions authors employ in visualization can significantly affect readers' interpretation. Kosara and Mackinlay [9] review the storytelling literature, present storytelling examples and scenarios, and provide an outline of the promising research directions. These discussions indicate that storytelling is a promising field for visualization, inspiring VisJockey to improve the communicative power of visualization.

3. VISJOCKEY CONCEPT

Among the seven genres Segel and Heer identified with 58 narrative visualization, the "magazine style genre" is the most common genre for static visualizations. However, it has not been the case for interactive visualizations [14]. VisJockey is designed to bridge the gap between the author and readers of these magazine style articles (i.e., data stories) with interactive visualizations, improving their readability. To compensate the author's absence while readers are reading through an article, VisJockey encodes the author's intended view for an illustratable text segment as a VisJockey play, which can be preloaded and then run later by readers.

Each VisJockey play is defined as a combination of interactions the visualization already supports and visual effects VisJockey provides. For example, Figure 2 depicts the sequence of one VisJockey

VisJockey Play

(b) start → (c) explore → (d) inset graphic → (e) end

(a) trigger play on text

As a sign of how quickly the gains have piled up, the Dow started the year in the low 13000s. It catapulted over 14000 on Feb. 1 and then took just 95 days (66 trading days) to burst above 15000. After some choppiness through the summer, the Dow has risen six weeks in a row.

Figure 2: An example of VisJockey play from a reader's perspective. (a) First, a reader triggers VisJockey play linked on text segment. (b) VisJockey play starts by adjusting all parameters (e.g., year span) to its start position and remove all graphics generated from previous play. (c) The line chart animates transition to explore from 1970–2010 to 2012–2013. (d) As the animation ends, it adds inset graphic that describes the growth of Dow Jones Index score, and (d) ends the play.

play. A user first triggers this play by clicking on the text segment followed by a symbolic icon (i.e., 🎬) which hints some animation embedded in this play. The play starts by adjusting parameters to its corresponding start position and remove graphics from previous VisJockey plays. Visualization starts reconfiguring its year span from 1970–2014 to 2012–2014 to focus on the author's statement. In this process, we use a visual effect of animation to maintain reader's attention. As soon as the animation ends, visualization starts adding inset graphics which are data points and arrows connecting them. This inset graphic aims to show the author's main point: the index score increased from 13000 to 14000 and then 14000 to 15000. Readers can trigger the same action by clicking on the same item on VisJocky menu (i.e., play list).

Our goal is not to select the most suitable visualizations for text segments automatically but to enable authors to illustrate their intended visualizations they believe to be the most appropriate. The article creation process with VisJockey consists of three steps: 1) find text segments that can be illustrated with VisJockey; 2) design and preload VisJockey actions; and 3) design an appropriate interaction trigger on text (provide flow diagram).

To be linkable, a text segment needs to have a direct or indirect connection to a range of data points. Figure 3 shows an example of linkable text segments. The terms like *The Dow*, *recently*, and *rose 109 points* provide information about related data ranges, so they can be linkable text segments.

Once text segments and their corresponding data ranges are defined, we may design and preload visualization effects on the text. There are various ways to load such effects. For instance, we can change the visual layout, encoding schemes, and text labels to highlight relevant areas of interest. It is often desired to use animated transition to deliver visual effects because it helps maintain readers' attention and awareness on data space. Additionally, we use a variety of inset graphics (e.g., arrows) to provide more direct relevance to the text triggered by users (see Figure 4).

4. VISJOCKEY STORY EXAMPLES

In this section, we provide three examplary stories to demonstrate thhee applicability of VisJockey. Note that our purpose is not to explain the content of articles, but to show different types of plays

The Dow recently rose 109 points to 16009.9. The S&P 500 traded as high as 1802 Monday morning and recently hovered around the millennium marker.
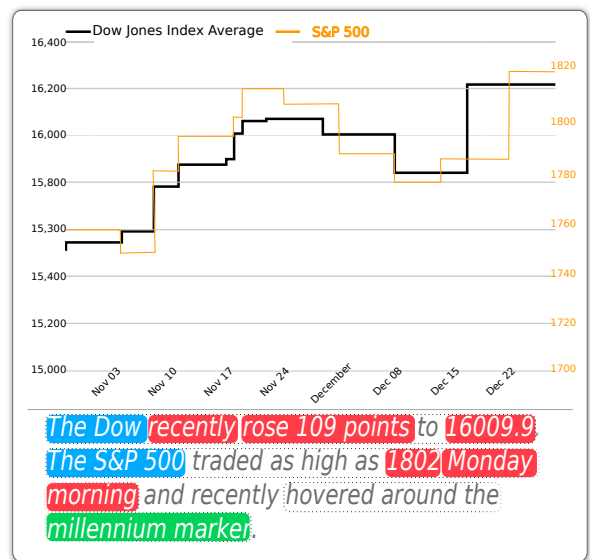
Figure 3: An article excerpt with the corresponding visualization. Candidates for linkable segments are framed by rectangles. The dotted rectangle indicate merged linkable segments.
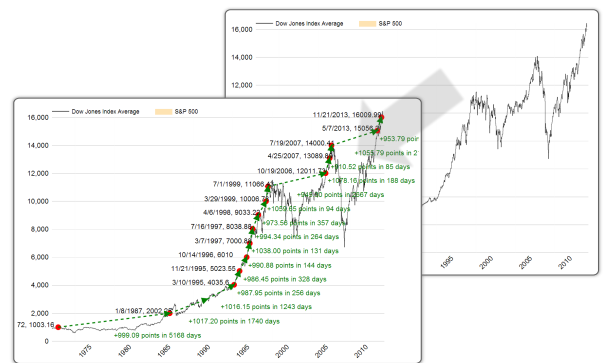

Figure 4: Application of inset graphics to a simple line chart. Interesting spots and trends of the chart are labeled with additional text, color overlays, and arrows.

in different visualizations. VisJocky is designed to share the stories on the web. Thus, we used html for layout, D3.js and Javascript for data handling, SVG for graphics, CSS for styling inline text, and jQuery for user interactions. Three of the authors implemented these examples. It approximately took two hours for a developer to create VisJockey plays for each story after text and visualization are implemented.

4.1 Dow Jones Index Score with Line Chart

We borrowed a news article on Dow Jones Index Score, describing the trend of scores over a period of time, from Wall Street Journal [13]. Since this original article only includes a table of data, we built a line chart (x-axis for time and y-axis for index score). This example includes ten plays that explore to specific time spans with animated transitions as shown in Figure 2. All text segments specify a range of years (e.g., in 2000s). Then, the line chart smoothly reconfigures its x-axis from the current range to the target range using animated transition. To help readers make the connection between the text and visualization, we provided animated visual effects that draw the data points (mentioned in the text) and arrows connecting the points, in a chronological order (Figure 3). If we provided a line chart ranging from 1900 to 2000 without any preloaded actions, it would be difficult for readers to find authors' intended data points and relate them back to the text content. With smooth transitions and text annotations, visualizations could switch their zoom levels and highlight corresponding data points related to the clicked text.

4.2 Automobile Story with Parallel Coordinates

In this example, we adapted a blog article "Parallel Coordinates," Robert Kosara published in his blog [8], describing how to read and interact with parallel coordinates. We also adapted interactive parallel coordinates made by Jason Davies from the D3.js library[1]. This example include 20 plays in total. The first five plays illustrate the transformation process from a table to a parallel coordinates visualization, as illustrated in Figure 5. This process consists of four phrases: 1) draw a table with the original data set; 2) detach cells and place them on parallel coordinates; 3) reorder the cells by sorting by corresponding axis; and 4) draw lines connecting cells; and 5) highlight one axis (i.e., Cylinders). We used animated transitions and inset graphics for the five plays to help readers understand what makes coordinates and lines in parallel coordinates more easily. We also had two text segments that show results highlighted by a range of manufactured years as Figure 6 shows.

---

[1]retrieved from `http://bl.ocks.org/jasondavies/1341281`

Figure 6: Parallel Coordinates. Automobiles from 1970 to 1982 with five core dimensions (i.e., mpg, cylinders, horsepower, weight, year) in corresponding axes. Correspoinding lines are highlighted as users click on a sentence about '1980 to 1982'.

4.3 The Wealth and Health of Nations with Bubble Chart

We created an article (Figure 1) from Hans Rosling's presentation on the wealth and health of nations, titled "200 Countries, 200 Years, 4 Minutes." We transcribed his narration from the video[2] and adapted a bubble chart visualization from D3[3]. We added introduction paragraph to show how to interpret visual encoding; we highlight one visual element at a time when a sentece explains it. This article also compares nations and geographic regions. We provided different color schemes to highlight the targets of comparison. Then, we used animation to play from a year to another. This highlight of colors can provide visual cues for attention. In addition, we provided additional inset graphics to illustrate indirect/fuzzy reference to visualizations. For instance, a part of article mentions "rich countries," which are hard to quickly grasp. By providing a dotted line across incomes higher than $10000, the reference becomes clearer for readers to understand.

---

[2]retrieved from `http://www.gapminder.org/videos/200-years-that-changed-the-world-bbc/`
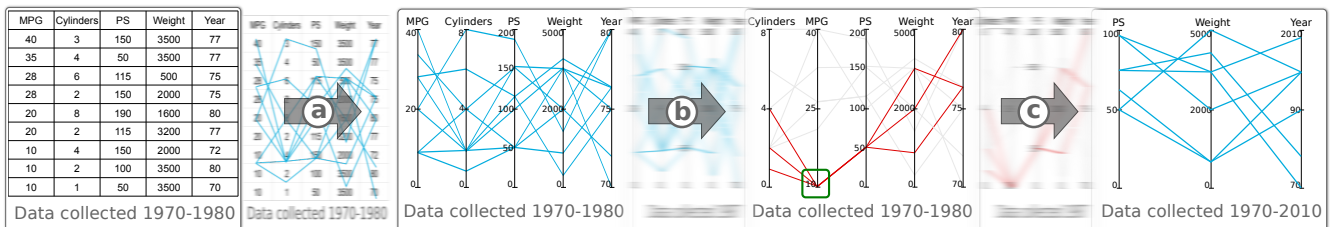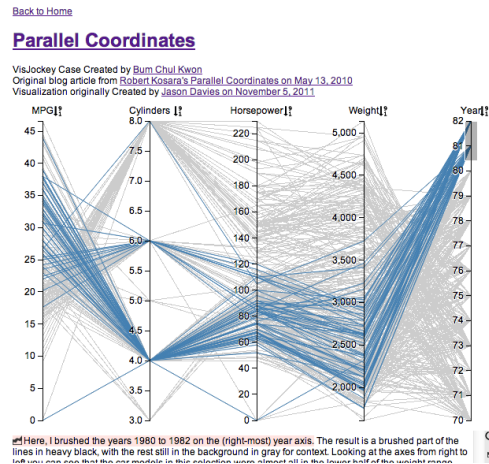[3]retrieved from `http://bost.ocks.org/mike/nations/`

Figure 5: Example instances of interactions, transitions, and inset graphics with parallel coordinates. (a) Visualization Change is applied to transform from table to Parallel Coordinates. (b) Reconfigure data dimensions by (Ordering), Encode new visual mappings by (Visualization Change), and highlight a subset of data by (Filtering). It also adds an inset graphic, rectangular box, on highlighted data point. (c) Explore to show more year and less dimensions by (View Transformation) and (Data Schema Change).

5. DISCUSSION AND CONCLUSION

VisJockey provides a new way to read stories connected with visualizations. Readers can read articles line by line and find relevant visualizations by interacting with the linked segments. This could offload cognitive burdens from readers, who are responsible to operate interactions to reach authors' intended view. Many current news articles often provide slide-show visualizations, which mimic the presentation slides, and infographic style visualizations to tell a story about data. These genres are very useful for some purposes, but they are limited to provide contextual information with straightforward prose as in traditional magazine-style articles. VisJockey can enrich magazine-style articles so that author-intended data stories can be unfolded on readers' interaction.

Our technique also encourages authors to use information visualizations to support their arguments and express complicated stories via VisJockey. Our application examples show how we can integrate visualizations within already-existing news articles. Creating VisJockey is not difficult given that D3.js is already widely used in news media, such as New York Times and Guardian. Here are some tips for article authors using VisJockey. One should note that VisJockey creation requires iterative tuning: writing an article, building visualization, linking text segments with visualization, and modifying text segments to clarify related data points. As we noted, our text segments sometimes miss references to exact data points. Without clarifying them, we may create misleading VisJockey plays. This implies that authors should refine their text to deliver messages more clearly with help of visualization. In addition, authors should fix a starting point of each play preferably on the previous play so that plays are sequenced naturally. During our development, we sometimes clicked through plays to fast-forward authors' points; we expect normal users would do the same thing occasionally. Thus, it is recommended to make VisJockey plays tell the story even by themselves.

As stated, building VisJockey plays require some programming efforts in addition to building visualization with D3.js. This creation process may be challenging for journalists, who may not have programming skills. This illuminates our future work. Our next goal aims to support VisJockey creation with a semi-automatic authoring tool. Our tool support will aim to incorporate various methods to select linkable text segments within a given article and combine them with intended visual effects. The tool should be designed with a graphical user interface, in which authors can directly manipulate objects to create their own visual effects and simulate plays at the same time.

VisJockey can play an important role to make articles more expressive by replicating the role of enthusiastic presenter. In oral presentations, we often easily understand what slides intend to show by associating them with presenter's talk. In current news visualizations, we observed that presenters are missing; in other words, authors have no method to associate their arguments with embedded visualizations. This study highlights the importance of unifying a message by creating clear links between text segments and visualizations. We cannot replace Hans Rosling's enthusiastic presentation skill, but we can provide a method to design the story to be read in author's way–with VisJockey.

6. REFERENCES

[1] Story points | tableau software. [Online; accessed 27-August-2014]: http://www.tableausoftware.com/learn/tutorials/on-demand/Story-Points.

[2] U. Brandes, B. Nick, B. Rockstroh, and A. Steffen. Gestaltlines. *Computer Graphics Forum*, 32(3pt2):171–180, 2013.

[3] N. Gershon and W. Page. What storytelling can do for information visualization. *Communication ACM*, 44(8):31–37, 2001.

[4] J. Heer, F. v. Ham, S. Carpendale, C. Weaver, and P. Isenberg. Creation and collaboration: Engaging new audiences for information visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, number 4950 in Lecture Notes in Computer Science, pages 92–133. Springer Berlin Heidelberg, 2008.

[5] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, Dec. 2011.

[6] J. Hullman, N. Diakopoulos, and E. Adar. Contextifier: Automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2707–2716, New York, NY, USA, 2013. ACM.

[7] N. Kong, M. A. Hearst, and M. Agrawala. Extracting references between text and charts via crowdsourcing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Canada, in press. ACM.

[8] R. Kosara. Parallel coordinates. [Online; accessed 19-March-2014]: http://eagereyes.org/techniques/parallel-coordinates, May 2010.

[9] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5):44–50, 2013.

[10] B. Lee, R. Kazi, and G. Smith. Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2416–2425, 2013.

[11] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, 2012.

[12] M. Rooze. News graphic collection. [Online; accessed 31-March-2014]: http://http://collection.marijerooze.nl/.

[13] S. Russolillo. Chart of the day: Dow tops 16000. [Online; accessed 19-March-2014]: http://blogs.wsj.com/moneybeat/2013/11/18/chart-of-the-day-dow-tops-16000/, November 2013.

[14] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 2010.

[15] F. Viegas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. ManyEyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007.

[16] W. Willett, J. Heer, J. Hellerstein, and M. Agrawala. CommentSpace: structured support for collaborative visual analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3131–3140, New York, NY, USA, 2011. ACM.