Computers & Graphics xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Computers & Graphics

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

Technical Section

An exploded view paradigm to disambiguate scatterplots☆

Q1 Salman Mahmood*, Klaus Mueller

Q2 *Stony Brook University, USA*

A R T I C L E I N F O

Article history:
Received 20 October 2017
Revised 26 February 2018
Accepted 26 February 2018
Available online xxx

Keywords:
Data transformation and representation
Zooming and navigation techniques
View-dependent visualization
High-dimensional data

A B S T R A C T

Small multiples is a popular visualization technique for dealing with overdraw in multi-class data. Small multiples are great at showing pieces of data individually, however, they do not explain how the different pieces fit together. They can also be difficult to understand for unacquainted users. We propose an interactive technique which uses the paradigm of exploded views to make small multiples visualizations more intelligible for unacquainted users. An exploded view is a drawing in which the different components of the object are separated by distance in such a way that the relationship between these components becomes apparent and hidden components of the data are revealed. We use the exploded view paradigm to create various animation designs for multi-class data. The designs are then compared using the Elo ranking scheme. We hypothesize that the exploded view animations increase the ability of users to appreciate the relations among data clusters (in the compound view) and at the same time get a clearer idea about the features of the individual data clusters (in the exploded view). We conduct a user study to compare this interactive approach with a compound view and an animated small multiples visualization.

1. Introduction

Scatterplots are a popular visualization technique for multi-class data, and due to their flexibility, they are used in a variety of contexts. They are, however, susceptible to the problem of screen clutter or overdraw. Overdraw is the conflict between screen space and large amounts of data. It can diminish the usefulness of a visualization by obscuring parts of the data, making it difficult to observe crucial characteristics of the data such as density, outliers, clusters, etc. The overdraw problem is amplified when dealing with multi-class data because when the point distributions of different classes overlap each other it is difficult to see the individual classes.

One of the most popular methods to reduce the effects of overdraw is small multiples. Small multiples is a clutter reduction method in which the data is broken down into multiple subsets and displayed separately, commonly in the form of a grid. There are various visual analysis tools that generate small multiples, such as ggplot2 library in R [1], the Polaris system [2] etc. Anand et al. [3] use a randomized non-parametric approach to partition variables and generate the most promising small multiples. Small multiples effectively divide the visualization into multiple parts, showing the features of the individual components (density, outliers, clusters etc). However, the relations between the different parts, such as overlap, distance between the clusters etc, are lost. Small multiples can also be confusing for users who are unacquainted with them. In this paper, we use the analogy of Exploded View diagram to make small multiples more intuitive for non-expert users. Ruchikacharon et al. [4] show that analogies can be an effective way to explain visualization techniques to users who are not familiar with them.

In an exploded view the components of an object are moved away from their original locations and suspended in the nearby space, giving the impression that the object is mid-way through an explosion. Exploded View is a very old concept; the first exploded view diagrams can be traced back to the fifteenth century [5]. Some of the earliest examples of exploded views diagrams were created by the famous Italian painter Leonardo da Vinci, who used the technique to show the inside of the human body or to show the internal mechanics of a machine. An example of Leonardo's exploded view can be seen in Fig. 1.

Using the exploded view paradigm to animate small multiples has two advantages. First, it makes small multiples more accessible to people who are not familiar with them. Even though the exploded view is a slightly more complicated structure, the average person has already been exposed to exploded views since they are commonly found in descriptive manuals of various do-it-yourself assembly equipment as well as LEGO manuals. Therefore, the user is already familiar with the paradigm of exploded

☆ This article was recommended for publication by X Yuan

* Corresponding author.
E-mail addresses: samahmood@cs.stonybrook.edu (S. Mahmood), mueller@cs.stonybrook.edu (K. Mueller).

https://doi.org/10.1016/j.cag.2018.02.008
0097-8493/© 2018 Elsevier Ltd. All rights reserved.

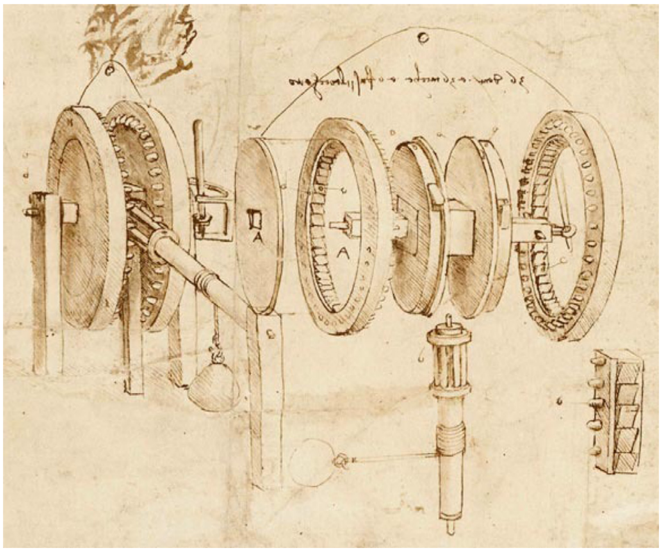

Fig. 1. Exploded view of a gear assembly, taken from Leonardo da Vinci's Codex Atlanticus. This is one of the earliest examples of an exploded view diagram.

views and will not be startled to see it. Second, while small multiples excel at showing the structure of an individual class (density, point distribution, outliers etc.), the relationships between the different classes (overlap, distance between subsets of data etc.) are not visible due to the spatial view separation. An exploded view is meant to show the individual components and the relationships between different components; therefore, by separating the points from different classes we can reduce the clutter and provide the user a better understanding of the relationships between the different classes.

We created multiple exploded view designs and then compared them using the Elo rating scheme to find the best of these. The Elo rating scheme was introduced as a chess rating system by Aprad Elo [6], and variants of this scheme are still in use to rate chess players. Elo uses pairwise comparisons (the chess match) to rate the different players. In our work, we extend its use to evolve exploded view designs and monitor their performance and ranking. After identifying the best exploded view design we compare it with both small multiples and compound views. Our study shows that exploded views are easier to understand and provide more information about the data.

Our paper is organized as follows. Related work is given in the next section. In Section 3 we explain how the various exploded view forces interact with each other and present several designs for exploded views. Two use cases of the different exploded view designs are presented in Section 4. In Section 5 we discuss the Elo rating scheme, and in Section 6 we present two user studies. Finally, we present some directions for future work and conclude the paper in Section 7.

2. Related work

In this section, we present previous work on overdraw reduction, the use of exploded views in visualization as well as the Elo rating scheme.

2.1. Overdraw reduction

Various techniques have been proposed in the literature on information visualization to reduce overdraw. Taxonomies and surveys have also been presented [7,8]. Ellis and Dix [9] analyze the advantages and disadvantages of different methods with the objective of creating a guide for matching different techniques to problems where different criteria may have different importance. The techniques suggested to reduce overdraw can be roughly divided into two parts: appearance-based and distance-based.

2.1.1. Appearance-based

This includes the type of methods that alter the appearance of the visualization in some way to cope with overdraw:

- *Size*: The size of the lines/dots can be changed, Woodruff et al. [10] use icons in less dense regions and small dots in dense regions.
- *Color*: Color blending and color weaving can be used to visualize multiple density fields. Chen et al. [11] use an algorithm to maximize the color distinguishability.
- *Opacity*: Opacity is useful in visualizing density as well as overlap in the data. Johansson et al. [12] show the utility of opacity in parallel coordinate plots.
- *Sampling*: If the amount of data available is too large, sampling can be used to reduce the number of data points. Here, density-based sampling techniques are able to reduce the risk of removing important data [13].
- *Filtering*: Filtering removes points that do not satisfy the criteria set by the user. Stone et al. [14] provide a window which can be moved around. The information inside the window is then filtered according to user specified criteria.

Some other methods in this category include aggregation [15], motion trails [16], blurriness [17].

2.1.2. Distortion based

Distortion-based methods are those that move the lines or points in the visualization in a way to reduces clutter.

- *Topological Distortion*: Topological distortion is a technique in which the topology of the plot is distorted using techniques like zooming, fisheye etc. Carpendale et al. [18] create a 3D surface and gives the user the tools to manipulate the surface of the data.
- *Displacement*: Points/lines are displaced to reduce clutter. Small multiples map points from different clusters into a matrix [19].
- *Dimension reordering*: This is generally used with parallel coordinate plots [20], where the order of the dimensions is changed in a such a way that the clutter is reduced [21].
- *Pixel Plotting*: Pixel plotting pack points onto a single pixel [22] or to empty nearby pixels to avoid overdraw.

This is by no means an exhaustive list and there might be other methods that are not mentioned here. It should be noted that these methods are not mutually exclusive, in fact, most visualization schemes will make use of more than one of these methods. Splatterplots [13] abstract information by grouping dense points into contours and sample the remaining points. They augment this with color blending to encode overlap between different classes. Color blending is less effective when there is overlap between many classes. Chen et al. [11] design a system that uses a hierarchical multi-class sampling technique that is augmented with dot-line representation for trend analysis. Both [11,13] are abstraction based techniques that may remove important details such as outliers and density. Conversely, the exploded view technique we present is a displacement technique. It can be used in conjunction with other overdraw reduction techniques. The choice of which techniques to use will depend on the data and the output the user is looking for.

2.2. Exploded view in visualization

Exploded Views have been used for visualization in a wide variety of ways. Li et al. [23] formulated an automated method

Please cite this article as: S. Mahmood, K. Mueller, An exploded view paradigm to disambiguate scatterplots, Computers & Graphics (2018), https://doi.org/10.1016/j.cag.2018.02.008