Towards Quantifying Multiple View Layouts in Visualisation as Seen from Research Publications
Al-Maneea, Hayder Mahdi Abdullah; Roberts, Jonathan C.

Published: 01/10/2019

Publisher's PDF, also known as Version of record

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Towards Quantifying Multiple View Layouts in Visualisation as Seen from Research Publications
Al-maneea, Hayder M.; Roberts, Jonathan C.

IEEE VIS 2019 Conference: InfoVis

DOI:

Published: 01/09/2019

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

Dyfnyiad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

17. Sep. 2019

Abstract:
We present initial results of a quantitative analysis of how developers layout the visualisations in their multiple view systems. Many developers create multiple view systems and the technique is commonly used by the visualisation community. Each visualisation shows data in a different way, and often user interaction is coordinated between the views. But it is not always clear to know how many views a developer should use, or what would be the best layout. We extract images of visualisation tools, across TVCG journal, conference, posters and workshop papers 2012-2018 to analyse the quantity and layout of the views in these visualisation systems. Focusing on view juxtaposition, we code the layout of 491 images and analyse view topology in juxtaposed views. Our analysis acts as a starting point to help designers create better visualisations, acts as a taxonomy of visualisation layouts, and provides a quantitative analysis of how many views developers have used in their visualisation systems.

Definitive Version at:
URL:

@article{Al-maneea-etal-IEEE VIS Conference-2019,
    author = {Al-maneea, Hayder M.; Roberts, Jonathan. C.},
    Conference = {IEEE VIS 2019 Conference-2019},
    title = {Towards Quantifying Multiple View Layouts in Visualisation as Seen from Research Publications},
    year = {2019},
    month = Oct,
    url = {},
    DOI ={}
}
Towards Quantifying Multiple View Layouts in Visualisation as Seen from Research Publications

Hayder M. Al-maneea*  
Bangor University and University of Basrah

Jonathan C. Roberts;† Member, IEEE  
Bangor University

ABSTRACT
We present initial results of a quantitative analysis of how developers layout the visualisations in their multiple view systems. Many developers create multiple view systems and the technique is commonly used by the visualisation community. Each visualisation shows data in a different way, and often user interaction is coordinated between the views. But it is not always clear to know how many views a developer should use, or what would be the best layout. We extract images of visualisation tools, across TVCG journal, conference, posters and workshop papers 2012-2018 to analyse the quantity and layout of the views in these visualisation systems. Focusing on view juxtaposition, we code the layout of 491 images and analyse view topology in juxtaposed views. Our analysis acts as a starting point to help designers create better visualisations, acts as a taxonomy of visualisation layouts, and provides a quantitative analysis of how many views developers have used in their visualisation systems.

Keywords: Information visualization, multiple view layouts

1 INTRODUCTION
Multiple view systems are often used by visualisation developers. But it is not easy for a developer to know how to layout and position the views in their systems, or how many views they should use, or what design attributes work best. We believe that developers and learners should have guidelines and frameworks to help them make good design decisions. Subsequently, we are keen to develop theories for visualisation, and specifically develop guidelines on best practices of view layout. But to achieve these goals, researchers need to perform basic research to understand best practices of what we currently do. The results of this paper helps us move forward towards our goals.

We present initial results of a quantitative analysis of the quantity of views used in multiple view systems as reported in the visualisation literature. In this paper we focus on two questions: (Q1) how many views, and (Q2) what are their arrangements? To answer these questions, we (1) prepare and extract images from papers, (2) code and classify each layout through visual inspection and discussion, we considered each visualisation in turn, judging the topological makeup of each visualisation, coded them such that we can classify them, and recorded a sketch of their topology and (3) analyse the results. Our three-stage methodology is shown in Figure 1, and we use this process to structure the rest of the paper. First the related work in Section 2. Second, we describe how we collected the images for our analysis (Section 3). Third, present how we codify the layouts (Section 4) and fourth we present the results for each question (Q1 in Section 5 and Q2 in Section 6 ) and discuss how we organised the sketches on a tabletop. Finally we discuss the results and their application, and conclude.

We extract images from research publications that were published at the IEEE Visualisation conference between 2012 and 2018. This seven year period provides a convenient and reproducible set of images of modern visualisation tools that have been designed and presented by community experts. In particular, because these works have gone through peer review, we assume that the authors have spent careful thought over how they present their tools, and consequently they have been attentive to the selection of their views and the presentation of their multiple view systems. We considered many sources, including using a general Internet search for visualisation images, video sources such as Vimeo or YouTube, or other online image repositories. Adding these sources might give us a rich data set of different images and this is certainly a limitation of what we did here. But, they also bring challenges, where image searches change over time, results change per user or geographic location, which would make it more difficult for others to confirm our studies, and to add more images for future years.

In this work we concentrate on view juxtaposition, where each views sits alongside each other, and on the topology of each design layout (e.g., a 2-view system can have one view above another, or left/right of each other). This short-paper extends our poster paper, that was presented at the 2018 IEEE Visualisation conference, where we introduced our methodology and highlighted initial results [1]. The feedback and discussions with the community at this event was invaluable; consequently we extended our quantitative analysis, widened our input data to include 2018 papers and included statistics on single-view systems. There are many research questions that we have investigated in our broader research project, including the quality of views, layout organisation, symmetry and design attributes, to the type of visualisation used. With space limitations of this format, therefore we summarise our main findings and focus on two main questions: Q1 How many views are used in multiple view systems? and Q2 What layout arrangements are popular in multiple view systems?

1 (1) Prepare
- Select files
- Extract images
- Name images

2 (2) Code
- Judge the views
- Sketch topology
- Code view type

3 (3) Analyse
- Layout on tabletop
- Analyse topology

Figure 1: (1) We extracted 491 images from IEEE VIS 2012–2018 conference publication. (2) We coded the images by their topology (making sketches of the layout, totalling 22 sheets of paper); discussing cases to confirm their layouts. (3) We cut the 22 sheets of sketches into individual tiles, and organised them on a tabletop, to analyse and tally the quantities.

2 RELATED WORK
We acknowledge the huge amount of well-cited research that has been achieved in the area of multiple views. Twenty five papers from the conference on Coordinated and Multiple Views conferences (2003 to 2007), rules and principles for the use of multiple views [24], state of the art in Coordinated Multiple Views (CMV) [20], juxtaposition, superposition and explicit designs for multiple view sys-
tems [8,9], keeping multiple views consistent [17], coordinating multiple views across large displays devices [12], and the many systems that create and use multiple coordinated views (e.g., ComVis [13], Snap-together [15], Waltz [19], Improvise [27], Jigsaw [22]), and the recent published work on the phraseology of multiple-views [21]. We acknowledge the huge quantity of research in the area of coordinated multiple view systems; where user interaction in one view is linked to another view (such as linked highlighting). It is through this linking that a user can better explore and discover interesting facts about the data. In this regard researchers have created rudiments of coordination [3], researched linked highlighting and linked navigation [15] and linked brushing [2]. In particular, developers have created many types of brushing including: compound brushing [5], multiple brushes [25], and complex filtering operations such as through angular brushing [10] or cross-filtered views [27]. But to date, there has been no systematic study investigating view-layout strategies for multiple views.

3 Preparation and image selection

To study the layout of multiple-view tools we needed a set of images to judge. Our goal was to capture images that demonstrate a visualisation tool or technique, that clearly originated from applications (whether on a desktop or website), that were created by through a snapshot/screen-grab operation or directly output from the tool.

Our preparation process had four stages. We started by considering all papers presented at the IEEE Visualisation conference between 2012 and 2018. But this creates a very large corpus of information that we decided to reduce the quantity, as follows. (1) We removed all files of supplementary materials. (2) We removed papers that did not have visualisations. (3) We removed papers that only had images that were clearly put-together or had been edited (by an image processing tool). We took this decision because it is difficult to quantify how much editing had been achieved by the authors. (4) We kept files that had at least one candidate image. We considered how we would code the images, whether manual or automatic encoding. We decided to use a manual coding process. We could have coded the layouts using an image processing algorithm, but decided that we would start with the manual encoding process. Therefore, we extracted the images from the files, and stored them in year-based folders, and labelled each image with a unique abbreviation (that we also use in LaTeX to cite the paper), as follows: Author1—Author2—Author3—ETAL—Year.png. This meant that we could easily reference the image, and locate the associated publication. If there were several suitable images which were different then we collected both, and added a F (for Figure) followed by the number, to the file name (—F1.png, —F2.png, etc.).

4 Coding the layouts

To address our two questions (Q1 and Q2) we needed to first be able to identify individual facets of a multiple-view display, and then code them effectively. In this section we define our view identification and coding strategies.

4.1 View identification

It is not necessarily easy to identify, or count, views on a visualisation. Sometimes it is clear that there are separate dividing parts to the view, that (say) one visualisation is a scatterplot and another a bar chart. But other times it is less clear how many sub component parts the visualisation contains. Designers overlay visualisations, or even place visualisations in an irregular way which can complicate deciphering the layout. But, because developers want users to understand their data display, they are deliberate in how they position their views. In addition, they will engineer the display such that the graphical marks standout from the background, and contain clearly perceptible parts. In perceptual terms, this is figure and ground [26].

The graphical marks that encode the data are the figure, which stands out from the background colour.

Most visualisation tools have interface components, and may include menus, buttons, slider bars, legends, colourmaps, legends, etc. Sometimes these are integrated with a view, sometimes they are shown in their own window. In most cases it is possible to ignore these menus. However sometimes the menus take up a significant space. We could code these views as “menu”, however this would not allow us to be consistent in our coding, because other views have this interaction/menu integrated into the views. We could code them as “null” views (or information panels), but again this may skew the results. Consequently, we make a multi-criteria solution. If the menu is on the side, or along the top, and can be easily ignored without changing the topology of the view layout we ignore it! If a menu-window is enclosed between other views then we treat this as “null” space and merge it into the closest neighbouring view. This allows us to treat all menus as part of (at least) one view, and every visualisation is treated consistently.

To perform meaningful and consistent manual coding of view quantity we needed to develop a set of rules (that we name the view identification process). We identified five rules to determine ‘what is a view’: (1) Views are usually visually separate from another view. Count the views that are clearly separated by spacing, a gap that is coloured in the background colour, rendered in a rectangle, or placed within a window. (2) Views have different tasks. Count the views separately if they afford a different task. (3) If we can name them, we have different views. For example, you could say “scatterplot, line graph and bar chart” and you would have three views. (4) Consider how a programmer would code it. If they cannot be separated visually, they may be able to be separated functionally. This is separation of concerns at the functional level, e.g., draw.scatterplot(). (5) Ignore interface components. Ignore menu windows if it is sensible to do so (such as a menu along the top of all views). We printed these rules, and kept them close, and especially used them when we had a dilemma of how to judge the layout.

4.2 Coding the layout arrangements

Now, being able to identify a “view” we needed a way to record the views and their configuration. We chose an inductive strategy to develop codes to answer Q1 and Q2. We developed these codes through refinement and critical thought. We chose a visual coding scheme where we displayed each image, made a simple sketch of the layout topology and labelled the sketch with the file name. If we had a dilemma on how to sketch the topology we drew all possible arrangements, which were later discussed. This meant that the ‘codes’ were developed as part of the judging process. Our coding-layout process had eight stages: (1) Every image was displayed on a computer screen. (2) Judge topology (using the view rules in view-identification process). (3) Code the layout in sketches, write file name alongside. (4) With dilemmas, sketch all possible layout configurations. (5) Discuss each dilemma, agree on one topology, and keep agreed topology. (6) Cut the sketches into individual tiles. (7) Arrange tiles on tabletop, discuss ideas, and categorise layouts. (8) Record quantities.

When we evaluated all 491 images, we generated over 22 sheets of paper (with on average over 22 sketches per page). We discussed 124 cases and agreed their topological structure, and randomly checked another 10 to make sure we agreed with a selection of the other cases. The agreed sketches were then cut up into individual tiles, keeping only the agreed topologies.

There are different potential ways to name the view layouts. We could use real names, such as “one view”, “dual view” or “three view”, but decided that this would not make a convenient shorthand version, because ordering of them would be difficult. Our solution was to label them with a number (the view quantity) followed by...
It is interesting that we chose the same label, in this case (two irregular layouts with little structure) is a small gridded layout.

We recorded the quantity of views used per year, and tallied each of these systems. We used a tabletop strategy to confirm the quantity of layouts. We arranged cut tiles on the table, organised by (first) the quantity of views, and then by their topological arrangement. We did this by arranging the cut tiles into groups, organised by their quantity; shown in Figure 1(right). By physically moving these tiles, it helped us better understand the frequency of each layout. We grouped tiles together by quantity of views, and then by their topological structure. E.g., putting all dual-views, three-views, four-views together, and so on. This tabletop view gives a physical area chart of the quantities in each strategy and from this tabletop collection of tiles we were able to quickly record the quantities in a spreadsheet for further analysis. We recorded the quantity of views used per year, and tallied each of the different view topologies up to 20 views. We decided that the space allocated to individual views, when the view quantity exceeds 20, views is very small, and therefore we bin all this information into a bin 20+. In the Tabletop view (Figure 1) we put them together under a label of “lots”. We tallied the scores and charted the data, shown in Figure 2.

| Table 1: Numlacature for the layouts. Those that are not similar to others are binned together into the category X. |
|---|---|---|---|---|---|---|
| 1A | 4A | 5A | X | 7X | |
| 2A | 4B | 5B | 8A | X | |
| 2B | 4C | 5C | 8X | X | |
| 3A | 4D | 5D | 9A | X | |
| 3B | 4E | 5X | 9X | X | |
| 3C | 4F | 6A | 10X | X | |
| 3D | 4G | 6B | 11B | X | |
| 4H | 6C | 11X | X | |
| 4I | 6X | |

A letter (indexing a view layout). For example, “1A” is a one view (there is only one possible layout), and “2A” is a vertical 2-view layout, while “2B” the horizontal layout. While the number is logical, it is unclear how to allocate letters to view configurations. Our solution is arbitrary. It was decided on a first-come basis. When counting the views, the first new type was allocated an A, the next unique type a B, and so on. Our naming scheme is listed in Table 1. Additionally, when the views quantities become large, there are many more topological arrangements. We could name every configuration, but decided that it would not be useful for our analysis, therefore we named these as being irregular and gave them the letter “X”.

5 RESULTS AND DISCUSSION: Q1, HOW MANY VIEWS?

We used a tabletop strategy to confirm the quantity of layouts. We arranged the cut tiles on the table organised by (first) the quantity of views, and then by their topological arrangement. We did this by arranging the cut tiles into groups, organised by their quantity; shown in Figure 1(right). By physically moving these tiles, it helped us better understand the frequency of each layout. We grouped tiles together by quantity of views, and then by their topological structure. E.g., putting all dual-views, three-views, four-views together, and so on. This tabletop view gives a physical area chart of the quantities in each strategy and from this tabletop collection of tiles we were able to quickly record the quantities in a spreadsheet for further analysis. We recorded the quantity of views used per year, and tallied each of the different view topologies up to 20 views. We decided that the space allocated to individual views, when the view quantity exceeds 20, views is very small, and therefore we bin all this information into a bin 20+. In the Tabletop view (Figure 1) we put them together under a label of “lots”. We tallied the scores and charted the data, shown in Figure 2.

From our analysis, we are now able to start to answer Q1: “how many views do people use”, and because there are different ways to interpret the statistics, we break this question into several further questions.

What is the most common layout? To answer this question we rank the view quantities (Figure 2). The most common layout is a 3-view system. Four-view systems are next, followed by one-view systems and dual-views. Six view systems are slightly more frequent than five view systems. There is a clear division between 6-view and 7-view (and more) systems. In fact 84.68% of the systems are 6-view or less. These results are important. They say that the majority of developers use 6-views or less, and most of them choose a 3-view system. The also suggest that (in general) fewer views are used more often, which supports the rule of parsimony by Baldonado et al. [24].

What is the average quantity of views used? The naive arithmetic mean calculates to 4.9, but this is misleading. We have a positive skew in the distribution of the view count (skew is 1.049), and it is clear from Figure 2 that we have a very long tail. Such a positive skew is understandable; when we count views it is impossible to get a value less than a 1-view system, and it is far less likely to see systems with huge quantities of views (it is just impractical to have a system with hundreds of views). We can demonstrate this situation by modelling a normal distribution from 1 to 20, with an average of 3 (as per the most frequent occurrence), and comparing our observations with this model. In fact we get statistically similar results to our coding observations: a Pearson correlation is calculated as 0.960 with a test p(0.885).

6 RESULTS AND DISCUSSION: Q2, VIEW ARRANGEMENTS

From our tabletop layout of the tiles, we were able to examine the view layouts. We exchanged the tile sketches into the nomenclature, as explained in Section 4.2 and shown in Table 1. Results for the top ten views are shown in Table 2. We have complete data for all view counts up to 20, but only show 10 views due to space limitations in this paper. The results, provide even more fine grain detail. To calculate these results we performed some vertical aggregation. Consequently if the topology was the same on the left (such as 3A) and to the right (such as 3B) we chose the same label, in this case 3A. While we miss out on calculating if the views are more left biased or right, it is a pragmatic decision that allows us to simplify the tallying.

Analysing the layouts in fine detail is interesting but challenging. There are many layouts, and as the quantity of views increase so does the number of arrangements. From Table 2 we notice the higher ranking layouts such as 2A, 3A, 3B. It is interesting that the popular layouts are not necessarily those with fewer views, for instance 5X and 6X (two irregular layouts with little structure) are within the top ten. Layout 6A is a small gridded layout and features also with the top ten. We also see that there are many familiar structures. 44 views show side-by-side views, and a further 18 have a two-way split (top to bottom). Four layouts have 3-views, and nine layout strategies with 4-views. We notices a two-thirds layout, while “2B” the horizontal layout. While the number is logical, it is unclear how to allocate letters to view configurations. Our solution is arbitrary. It was decided on a first-come basis. When counting the views, the first new type was allocated an A, the next unique type a B, and so on. Our naming scheme is listed in Table 1. Additionally, when the views quantities become large, there are many more topological arrangements. We could name every configuration, but decided that it would not be useful for our analysis, therefore we named these as being irregular and gave them the letter “X”.

For our analysis, we are now able to start to answer Q1: “how many views do people use”, and because there are different ways to interpret the statistics, we break this question into several further questions.

What is the most common layout? To answer this question we rank the view quantities (Figure 2). The most common layout is a 3-view system. Four-view systems are next, followed by one-view systems and dual-views. Six view systems are slightly more frequent than five view systems. There is a clear division between 6-view and 7-view (and more) systems. In fact 84.68% of the systems are 6-view or less. These results are important. They say that the majority of developers use 6-views or less, and most of them choose a 3-view system. The also suggest that (in general) fewer views are used more often, which supports the rule of parsimony by Baldonado et al. [24].
weight of both sides. We propose that more designs are symmetrical than not. To explore this hypothesis we plot the quantity of symmetrical views (whether vertical or horizontal symmetry) against the non symmetrical. We do observe that there may be a trend to utilise more symmetrical layout strategies. But our results are not conclusive, because we also see that there is a natural tendency for layouts with more views to be less symmetrical. And it is difficult to tease apart these two observations. Another observation is that the thin views (such as 3D $\ddagger$, 4G $\ddagger$, 4H $\ddagger$) are lower in the rankings. These views seem to be used for timelines, or line graphs, would be placed in long and thin layouts. Additionally, forms, it would seem sensible, from a design standpoint, that there are options for developers:

- Developers of multiple view systems have used many different layout strategies to position their views on the screen. But there is commonality in design choices, and we have discovered that 3-view layouts are the most popular, and that bar charts and scatter plots are widespread, and that together with line charts, heatmaps and node-link diagrams they form over 50% of all views in multiple view systems. We have discovered that the top eight visualisation layout strategies are 1A $\ddagger$, 2A $\ddagger$, 3A $\ddagger$, 3B $\ddagger$, 4A $\ddagger$, 5X $\ddagger$, 3A $\ddagger$, 2A $\ddagger$ (see Table 2), and that there appears to be preference for symmetrical layouts. We have demonstrated that although there is a preference for fewer views, some developers still display systems with a huge quantity of views (layouts with 7-views and more equals over 15% of the total quantity of layouts). But we are sure that there is still much information hidden in this data that could be extracted. In particular we hope that our results and analysis will restart the debate over how many views are suitable, and will help to focus the minds of the developer as they create multiple view systems, to contemplate how they are laying out their views, and how many views they are using.

One aspect, in particular, that we feel should be debated and further researched, is the question of who has control over the layout? If we look programming languages, we find structures such as the GridBagLayout in Java, to help programmers layout structures. We see panel layouts in web structures (such as top, left, main, right or bottom panel), and templates to help Web developers layout their information. But there has been little research in the best layout strategies for visualisation. Certainly it will be developer of a visualisation system who will determine how much control the user has over the tool, and over the layout of views, and we propose five options for developers:

- Developers can predefine the layout. This is a fixed strategy, and is usually used reserved for fewer views or bespoke systems (designed for a particular purpose or user). For instance, in two view systems there is little choice: the views can be laid out left/right or above/below. Note such systems are also known as side-by-side, parallel or dual view [11, 14, 16] systems, or if one view is more important then primary/secondary, focus+context or overview and detail systems.

- Views positioned based on data. For instance, Roberts [18] positions the views based on hierarchical tree of data exploration, or the splom layout [4] (lattice charts) positions the small multiples to allow pairwise comparison of scatterplots, and Polaris [23] and spreadsheet visualisation approach by Chi et al. [6] display information in grid-based layouts.

- Group views that are coordinated together. For instance, views that sharing linked highlight, or linked navigation can be positioned closely and maybe in the same window. I.e., the type of coordinated manipulation can be used to control the positions of views; Roberts named these "render groups", [18] and Weaver puts them side-by-side and visually connects them with lines and arrows [27].

- The screen size can be used to determine the layout: where for instance, small multiples are laid on the screen in order that wraps onto the next line, and as the window size is changed so the viewed quantity changes (in the same way that a responsive/mobile-aware adjusts the content determined by the width).

- User can determine the layout. Often systems are created whereby the user can drag (from a toolbox of possible visualisation types) and drop the selection onto a canvas, where the views are snapped to align together (such as with Improvise [27], Jigsaw [22], Vinca [7] and many other tools).

There are many more questions that we could ask, and there is a much further work to be done. E.g., while we have started to investigate the connection between view layouts and the visualisation forms, it would seem sensible, from a design standpoint, that there is a strong correlation between view type and position in the layout. With (for instance) long and thin structures, such as timelines or line graphs, would be placed in long and thin layouts. Additionally, there may be a connection between the view layout strategy and its position of the layout in the article. Where, for instance, visualisations that have more views are placed along the top of the article, with those with less views in a column. We leave these questions to future work.
REFERENCES


