Exploring Visual Information Flows in Infographics

Min Lu[1], Chufeng Wang[1], Joel Lanir[2], Nanxuan Zhao[3,4], Hanspeter Pfister[3],
Daniel Cohen-Or[1,5] and Hui Huang[1]

[1]Shenzhen University [2]University of Haifa [3]Harvard University [4]City University of Hong Kong [5]Tel Aviv University

{minlu, wangchufeng2018, huihuang}@szu.edu.cn, ylanir@is.haifa.ac.il, nanxuzhao2-c@my.cityu.edu.hk,
pfister@g.harvard.edu, dcor@tau.ac.il

ABSTRACT

Infographics are engaging visual representations that tell an informative story using a fusion of data and graphical elements. The large variety of infographic design poses a challenge for their high-level analysis. We use the concept of *Visual Information Flow* (VIF), which is the underlying semantic structure that links graphical elements to convey the information and story to the user. To explore VIF, we collected a repository of over 13K infographics. We use a deep neural network to identify visual elements related to information, agnostic to their various artistic appearances. We construct the VIF by automatically chaining these visual elements together based on Gestalt principles. Using this analysis, we characterize the VIF design space by a taxonomy of 12 different design patterns. Exploring in a real-world infographic dataset, we discuss the design space and potentials of VIF in light of this taxonomy.

Author Keywords

Infographics; visual information flow; design analysis.

CCS Concepts

•**Human-centered computing** → **Visualization design and evaluation methods;**

INTRODUCTION

Infographics are visual representations consisting of graphical elements and data components designed to convey an informative narrative. The way they combine visual elements and text into an organic story is essential to effectively convey their message. Unlike other visual media, such as interactive story boards [26] or data-driven video [1], infographics use *static* graphical elements, text, and notable embellishments, designed to help readers easily interpret the story. Understanding how these elements can be combined effectively can help create better infographics designs as well as guide the general visual organization of a story.

The design of infographics encompasses various creative means, making analysis of their visual design space difficult, mainly due to two aspects. First, infographics are composed

Figure 1. Examples from the analysis of our infographic repository, with the extracted visual information flows shown on the right of each infographic.

of various visual elements with diverse appearances, such as icons, images, embellishments, or text. Skillful graphic designers usually create them with an aesthetic and creative mindset, often injecting them with personality and style (e.g., cute, powerful, or romantic style) to achieve a certain atmosphere. Artists may also distort certain design elements (e.g., exaggerating the theme figures) to emphasize them. Second, the spatial arrangement of these visual elements is carefully chosen to convey a unique idea to the audience. Therefore, the arrangements are generally diverse, and do not necessarily follow a well-known structure.

In this work, we introduce the concept of *Visual Information Flow* (VIF), *the underlying semantic structure that links the graphical data elements to convey the information and story to the user*, as a means to understand visual organization of stories. We explore VIF in a broad range of infographics, with the goal of understanding the design space and common patterns of information flow, and ultimately supporting better infographics design, especially for novices. To tackle the challenges of diverse designs and arrangements in infographics, we leverage advances in automated image understanding. We collected a repository of around 13K design-based infographics and trained a neural network to locate the visual elements. We automatically construct the VIF from these visual elements using Gestalt principles that are often used by designers for effective visual communication.

Our method is able to extract VIF from a wide variety of infographics designs with various artistic decorations. Using the extracted VIFs of thousands of infographics, we are able to characterize the design space and present a taxonomy of VIF patterns as well as explore and analyze the collected infographics from different perspectives. As such, our main contributions are as follows:

- A method for the breakdown of infographics and the construction of its VIF from automatically detected elements

- A taxonomy of main VIF design patterns and exploration of the VIF design space

- A system that supports infographic search according to VIF patterns

- A large dataset of 13,245 infographic templates, out of which 4300 include annotated boundary boxes of elements, and a dataset of 965 infographics with real data.

RELATED WORK

Recently, several works have introduced data-driven analysis of large-scale datasets as a way to explore design spaces. For example, using a repository of existing webpages as input, Kumar et al. [25] explore their rendering styles. Liu et al. [27] mine UI design patterns via code-and vision-based analysis in mobile applications. In our work, we learn and analyze the VIF of a large number of infographics in a data-driven manner. Our work shares some similarity with emerging research on data-driven design and story telling in the field of HCI.

Infographics. Infographics have gained increasing interest among researchers recently. Bateman [3] showed that embellished charts do not reduce accuracy but increase memorability compared to plain charts. Haroz [17] examined how simple pictographic embellishments affect viewer memory, speed, and engagement within simple visualizations. Borkin et al. [5] conducted experiments to understand what makes infographics effective and memorable. Harrison et al. [18] showed that people can form a reliable first impression of an infographic in a relatively short time, and then studied the different design factors that influence this first impression.

Since the advent of deep neural networks (DNN), more and more works built on their strength to learn from large scale data. Bylinskii et al. [9] developed a model to extract textual and visual elements from an infographic that is representative of its content. Using a large collection of crowd-sourced clicks and importance annotations [23] on hundreds of designs, Bylinskii et al. [10] built a model to predict the visual importance of graphical elements in data visualizations. Zhao et al. [38] proposed a deep ranking method to understand the personality of graphic design directly from online web repositories.

In our work, we take advantage of DNN, Gestalt principles, and a large collection of data to study another essential aspect of infographics, visual information flows. Unlike previous work that tracked and studied users' reactions to infographics, such as attention [23] and memorability [5], our exploration of VIF aims to learn the design patterns from the perspective of infographics creators rather than end users.

Visual Information Organization. The importance of a good organization of information, visual information in our case, in story-telling is well recognized [6, 7]. Design patterns of organizing visual information have been studied in various forms. Siegel and Heer [34] analyzed 58 narrative visualizations and characterized seven distinct genres of narrative visualization, including comics, animation, slide show, and more. Via a survey of 263 timelines, Brehmer et al. [8] revisited the design of timeline in story telling and identified 14 design choices characterized by three dimensions: representation, scale, and layout. Looking at comics, Bach et al. [2] performed a structural analysis of comics and introduced design-patterns for rapid storyboarding of data comics. For slide shows, Hullman et al. [21] performed an analysis of 42 slideshow-style narrative visualizations and studied how sequencing choices affect narrative visualization. Animation is one of the main storytelling genres (which include video and movies) which can be effectively used to show narration [20]. Heer et al. [19] examined the effectiveness of animating transition to tell a story. In interactive visual narration, McKenna et al. [30] performed a crowdsourced study with 240 participants to examine the factors that shape the dynamic information flow with users' interactive input.

Lately, several visual authoring tools have been emerging to facilitate the visual organization of information and improve visual expressiveness in story-telling. Kim et al. [24] proposed a method for designing graphical elements enhanced with data. Continuing this line, Liu et al. [29] provided a systematic framework for augmenting graphics with data, in which designers draw vector graphics with familiar tools and then bind the graphics with data. Wang et al. [37] presented a visual design tool for easily creating design-driven infographics. Ellipsis [33] provided a user interface that allows users to build visualization scenes that include annotations in order to tell a story. Several timeline-based story authoring tools were developed (e.g., DataClips [1] and TimeLineCurator [15]). Recently, Chen et al. [11] developed a method to parse static timeline visualization images using a deep-learning model to enable further editing.

Most of these previous works focus on visual information design in *dynamic* media and *interactive* visualizations. Our work explores visual information flows in *static* infographics without user interactions, from which the distilled design patterns can empower the design of infographics authoring tools.

METHOD

The following section describes the methodology we used for automatic construction and analysis of VIF in infographics.

Overview

Infographics are a composition of *graphical data* and *artistic elements*, where the former convey the information, and the latter make the infographic visually appealing. An effective infographic is self-contained, meaning the whole information is contained in one image and conveyed to the user via VIF that connects the visual elements. Often, pieces of information are organized into *visual groups*, which are compound graphical

data elements for multi-facet information, e.g., an icon, a subtitle, or a textbox. A visual group usually illustrates the content using symbolic graphical depictions as well as textual details (Figure 2).
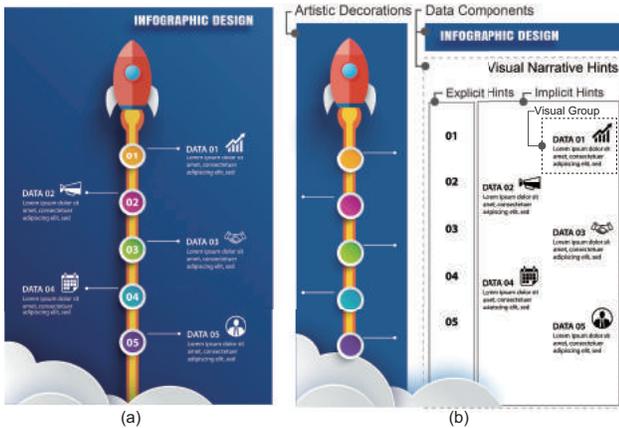

Figure 2. Infographics Model: (a) an infographic example. (b) separating the artistic decorations, the visual information flow of the infographic connects the visual groups of data elements in narrative order, hinted by explicit graphics (e.g., digits here) and implicit Gestalt principles.

The variety of artistic elements is rich, and the visual groups turn out to have very different styles, even within a single infographic. They may use different icons, color palettes, font families, graphics, and texts. To ease the parsing and the interpretation of the data, designers typically inject *visual narrative hints* into the infographics to guide the readers to effectively trace the information flow. There are two types of hints, explicit and implicit. Explicit hints use graphical data elements that suggest and index the flow, such as digits, arrows, or textual descriptions. Implicit hints come from various principles that designers follow to achieve a cohesive design [35].

Many of these design principles, such as unity, balance, or contrast, are less relevant to the narrative structure. On the other hand, the Gestalt principles of visual grouping perception [13] provide effective guidance on visual group identification and connection, and are very applicable in hinting the VIF. For example, visual elements placed close to each other are more likely to be a group (Gestalt Proximity Principle), and visual groups designed to look similar (Gestalt Similarity Principle) or placed to form an intuitively regular pattern (Gestalt Regularity Principle [36]) are more easily recognized. We use these Gestalt principles to automatically weave the extracted visual elements together into visual information flows.

We use a bottom-up methodology to automatically construct a VIF signature for a given infographic (Figure 3). The first step is to locate the visual data elements related to the visual information flow. Since there are various creative means to fuse data and graphics, it is impractical to extract elements based on heuristics alone. We use the power of machine learning and deep neural networks for image understanding to detect the data elements. Given a manually labeled training set, we train a neural network and identify the visual data elements

(described in detail in subsection *Data Element Extraction*). Then, we associate elements into visual groups based on the Gestalt proximity and similarity principles, and then trace various information flows among the visual groups based on the Gestalt regularity principle. The constructed flows in those trials are scored according to their regularity and the best one is picked as the visual information flow (see subsection *Information Flow Construction*). With the detected VIF, we then explore the VIF space to create a taxonomy of VIF design. Each VIF is associated with an icon image of uniform size, which serves as a VIF signature. Taking their VIF signatures as high dimensional features, infographics are embedded in a 2D space using t-SNE. Based on this embedding, a semi-automatic classification process is performed for the main design patterns (see subsection *Design Pattern Exploration*).

InfoVIF Dataset

Several studies created infographic datasets from visual content platforms, such as Flickr or Visually, for various purposes (e.g., [10] [32]). To the best of our knowledge, the single public available dataset related to infographics is MassVis[1]. This dataset is a collection of graphical designs from multiple sources, e.g., magazines, government reports, etc. However, many of the images in MassVis are highly specialized, e.g., illustrating scientific procedures or presenting statistical charts. In this work, we focus on more general infographics that are not customized for a particular range of subjects or domains.

We collected a large dataset of over 10K infographics, *InfoVIF*, from two design resources sites for graphics, Shutterstock[2] and Freepik[3]. The infographics collected in *InfoVIF* are mostly design templates aimed to be a starting point for domain-specific augmentations. This corpus was chosen for several reasons. First, it has a more uniformed style of visual elements. For example, design templates usually use the same size for each textbox, in contrast to end-infographics. This helps alleviate some of the technical challenges in the detection and construction process of the visual elements of the infographics as described later on. Second, the collected infographics are contributed by various world-wide designers, and are very diverse in their design themes and styles. Together, they provide a good coverage of the design space of infographics. Finally, the infographic templates are usually used as design resources from which people get inspired and adapt their own infographic design. *InfoVIF* potentially serves and represents the origin that establishes numerous end-infographics in various domains.

We gathered a broad range of infographics for our dataset, by searching the keyword 'infographics' in the two mentioned websites and pruned (i) the infographics composed of multiple subfigures; and (ii) those only with figures or texts. In the end, 13,245 infographics were collected in *InfoVIF*, 68% from Freepik and 32% from Shutterstock. InfoVIF is freely available for academic purpose at http://47.103.22.185:8089/.

---

[1]http://massvis.mit.edu/

[2]https://www.shutterstock.com/home
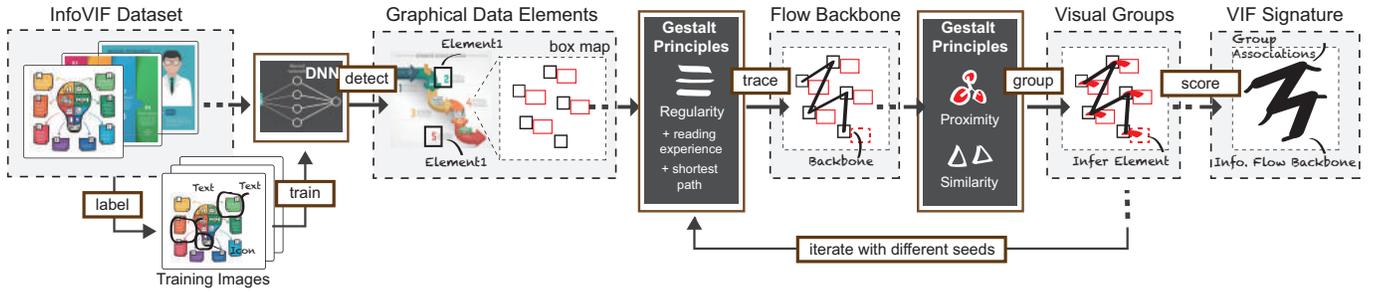
[3]https://www.freepik.com/home

Figure 3. VIF Construction Pipeline: using manually labeled infographics as training data, we use a deep neural network to detect visual data elements, leaving aside artistic elements. We run multiple passes to trace and score different information flows based on Gestalt principles, and then we pick the best one.

Data Element Extraction

For each infographic, the first step is to detach its graphical data elements from artistic decorations. This is a classic object detection problem. With the recent development of deep neural networks, object detection has benefited immensely by learning from large scale human-labeled datasets [16, 28]. We adopt YOLO [31], one of the state-of-the-art object detection methods, to solve the data element extraction problem.

Data elements are categorized into four main groups, *text*, *icon*, *index* and *arrow*. Text is further distinguished by two types, title and body text. For the index, we differentiate 18 numbers (1 to 9, 01 to 09) and seven main indexing letters (A to G). Arrows are discriminated in eight directions (e.g., left, left-top, top, etc.). It ends up with 36 labels for graphical data elements in total (see examples in Figure 4).

Figure 4. Output of Element Extraction: data elements are extracted at high precision (a, b), with (c) false positives, e.g., the home button, and (d) some misses, e.g., '$' icon.

Image Augmentation
To provide accurate training signals to the model, we randomly selected 4,300 infographics from *InfoVIF* and manually annotated them with the 36 labels as mentioned before. Finally

from the 4,300 infographics, we got a total of 61,848 bounding boxes and an average of 14 labels for each infographic. The annotation dataset is also released as public data resource at the link given before.

To alleviate the overfitting problem and help our model better generalize during testing time, we applied two data augmentation methods to the training dataset. The first is to convert images to gray scale. The second is cropping: we first rescaled the image from 100% to 200% with intervals of 5%, to generate a sequence of images with different sizes. For each size, we then cropped the region with the same area as the original image from four corners to produce different cropped images. We discarded the cropped image if it intersects any bounding box of the annotated elements. After augmentation, we ended up with 25k annotated training images.

Performance
Prior to training, we randomly selected 700 manually tagged infographics from the annotated training set and put them aside as the test dataset for performance evaluation. The model was trained for 70k steps using 25k infographics. We report performance using the precision-recall metric. We get a mean average recall of 0.75, and a mean average precision of 0.628 over the test dataset. Detailed precision numbers for different elements are shown in Table 1. The precision for Numbers and Texts are the best, reaching as high as 0.78. Figure 4(a) and (b) show some successful extractions of data elements. However, automatic methods like ours will always incur missed detection (e.g., the bottom icon in Figure 4(d)) or false positives (e.g., the button detected as 'icon' in Figure 4(c)). To aid in the construction of visual information flow, we use the Gestalt similarity among groups to infer the missing elements (to be introduced later on).

Class	Precision	Class	Precision
Number(1-9)	0.456	Number(01-09)	0.791
Body Text	0.782	Letter (A-G)	0.657
Icon	0.672	Title	0.695
Arrow	0.566	**Average**	0.628

Table 1. Precision of locating different data elements. The average recall for all classes is 0.75.

Information Flow Construction

Given the extracted data elements, we construct the visual information flow. This reverse engineering process is challenging since there are numerous possible solutions, and the sought pattern has no fixed spatial form (Figure 1). To identify the VIF, we rely on the Gestalt principles that implicitly guide infographic design. These principles (elaborated below) dictate the grouping of the elements and their relationships in the VIF. The basic idea is to trace the VIF by first forming the *flow backbone* (Figure 5(b)), and then expanding it by associating nearby elements as *visual groups* along the backbone (Figure 5(c)). We first select an element with high priority to form the seed of the backbone. Then we construct the flow backbone from the seed to other similar elements (Gestalt principle of similarity) tracing the shortest path. With the traced flow backbone, each of its elements is expanded to associate the elements in its neighborhood (Gestalt principle of proximity) to generate visual groups with similar configuration. We run this process on different seeds and score those flows according to the Gestalt Principle of regularity, finally picking the one with the highest score.
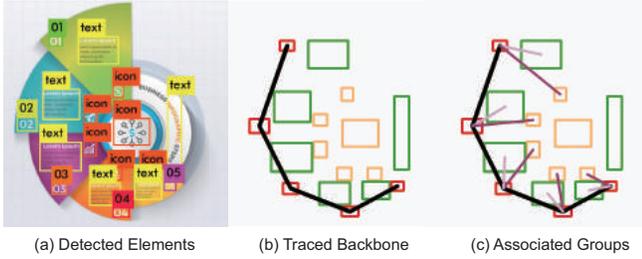

(a) Detected Elements   (b) Traced Backbone   (c) Associated Groups

Figure 5. Flow Construction: Using the detected data elements, we first trace the flow backbone and then associate nearby elements into visual groups.

Gestalt Principles in VIF

We identify three Gestalt principles that are fundamental to VIF: Proximity, similarity, and regularity.

Proximity within a group. Elements in a visual group are usually close to each other. For example, in Figure 4(a) and (b), the text and icon are nearby and naturally recognized as a group. This principle guides us to search for elements in the neighborhood when composing a visual group. The distance between elements can be considered using three perspectives: Euclidean distance, horizontal distance, and vertical distance. For example, in Figure 4(c), text and icon have the closest vertical distance, but are not close in Euclidean or horizontal distance.

Similarity among groups. Visual groups in an infographic are usually designed using similar visual configurations. Taking Figure 4(a) for example, the four visual groups are all composed of an icon and texts, though the icons and texts may not have the same design in different groups. This principle provides hints on how to grow the visual group and infer missing detected elements into a visual group. For example, in Figure 4(d), the bottom icon can be inferred with high probability by considering the existence of text on the right.

Regularity across groups. Visual groups in infographics are commonly designed with objects placed in structured, symmetrical, regular or generally speaking, harmonious patterns to achieve a pleasing and interesting visual effect. Conversely, infographics designers usually avoid crossing, or long distance jumps in the information flow. In the following section, we propose a set of measures to quantify the regularity of narrative flow by which we score the fitness of the information flow.

Flow Extraction

The flow construction procedure is described in the pseudo-code shown in Algorithm 1, which consists of five operations: (i) select seeds, (ii) trace flow backbone, (iii) compose visual groups, (iv) amend information flow, and (v) scoring the flows.

Algorithm 1 Flow extraction algorithm

```
 1: procedure EXTRACTFLOW
 2:     eleSet ← set of elements
 3:     seedList ← selectSeeds(eleSet)
 4:     flow ← [ ]
 5: top:
 6:     if seedList.len = 0 then return flow
 7:     seed ← seedList.pop
 8:     seedAllies ← seed + getSeedAllies(seed)
 9: loop:
10:     tempFlow ← traceFlow(seedAllies).
11:     vgroupList ← composeGroups(flow, seedAllies, eleSet)
12:     newAllies ← guessEles(vgroupList, flow, eleSet)
13:     if newAllies.len = 0 then
14:         flow ← scoreFlows(flow, tempFlow)
15:         goto top
16:     seedAllies ← seedAllies + newAllies
17:     goto loop
```

Select seeds. The *seed* is a selected data element from which we trace a tentative backbone. We select seeds with high potential in forming the information flow. We evaluate the detected data elements and assign them different priorities with the following criteria: index priority and shape similarity.

Index Priority. We prioritize elements that carry some semantics that suggest an indexing order, such as numbers and letters. Elements that contain text or icons get lower priority.

Shape Similarity. Visual elements with the same detected tag or shape similarity are considered *allies* of the seed. To avoid redundancy, we give low priority to elements that are similar to an existing seed. Shape similarity between element i and j is measured with:

$$Similarity\_element = 1 - max(|w_i - w_j|, |h_i - h_j|), \quad (1)$$

where w and h are the width and height, respectively, of the detected bounding box normalized to $[0, 1]$.

Trace flow. With the set of seeds and the set of similar elements (measured by *Similarity_element*), we construct a flow backbone by optimizing and trading off between the following three criteria: shortest path, regularity, and common reading order.

Shortest Path. Empirically, to achieve a clear and efficient visual communication, designers usually prefer to steer the information flow using the shorter distances between elements, to help the eye to naturally follow the elements.

Regularity. Elements are arranged in well-organized structures, e.g., with consistent spacing, in a symmetric or Euclidean geometric layout. In this work, we use regularity as a primary clue in tracing the flow backbone. Given a flow as a list of points $< p_0, p_1, ..., p_n >, p_i = (x_i, y_i)$, we evaluate regularity by the standard deviations (noted as S) of four sets: line segment lengths ($r_1 = \{|p_{i+1} - p_i|\}$), adjacent horizontal shifts ($r_2 = \{|x_{i+1} - x_i|\}$), adjacent vertical shifts ($r_3 = \{|y_{i+1} - y_i|\}$), and turning angles ($r_4 = \{arc(p_{i+1}p_{i+2}, p_ip_{i+1})\}$). The overall regularity of a flow is taken as the one with best regularity score among the four using:

$$Regularity = 1 - min[S(r_1), S(r_2), S(r_3), S(r_4)]. \quad (2)$$

Common Reading Order. Depending on the context, there are preferred high-level reading orders, e.g., from left or right, clockwise or counterclockwise. In this work, we take the most common reading order to decide the flow if no explicit hints exist, i.e., left to right horizontally, top to bottom vertically, and clockwise in case the elements have a radial arrangement.

Compose visual groups. Visual groups grow from the data elements that are chained along the backbone using the following expanding rules.

Elements in Proximity. Elements of a group are normally placed close to each other. Given an element on the backbone (denoted by (x_m, y_m)), we search for the elements in its three principal neighbors with priority, according to basic priorities of the backbone's shape. That is, elements in vertical neighborhood (i.e., $[(x_m - \delta_x, *), (x_i + \delta_x, *)]$) are searched first when the backbone is oriented horizontally (i.e., when the standard deviation of y-positions of elements E on the backbone is small enough) or vice versa. Specifically:

$$Proximity = \begin{cases} [(x_m - \delta_x, *), (x_i + \delta_x, *)], & S(E_y) < \theta_y \\ [(*, y_m - \delta_y), (*, y_i + \delta_y)], & S(E_x) < \theta_x \\ [(x_m, y_m) - (\delta_x, \delta_y), (x_i, y_i) + (\delta_x, \delta_y)], & others \end{cases} \quad (3)$$

.

Similarity among Groups. As discussed earlier, visual groups in infographics are usually designed with similar configuration (Figure 6). We measure the similarity between visual group i and j with the Jaccard coefficient:

$$Similarity\_group = \frac{|E_i \cap E_j|}{|E_i \cup E_j|}, \quad (4)$$

where E is the set of elements in the group, and two elements are counted as equal when they are similar (Equation 1).

Amend information flow. Missing elements are inferred by considering the remaining elements' affinities with other composed visual groups. For example, if an element is missing in a group, we can infer the element by cross verifying in

other groups. We place the missing elements according to the average placement of their counterparts in other groups.
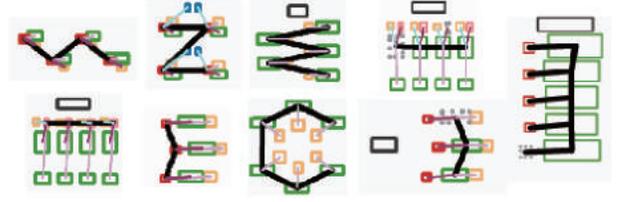

Figure 6. Constructed Flow Examples: The flow backbone (black polyline) is traced, and visual groups (with linked lines) are associated based on Gestalt principles, including amended elements (dashed rectangles).

Parameters and Performance
In the flow construction we use a set of parameters that we empirically fine tune, where the detected boxes are normalized to a canvas unit size. Two elements are considered to be similar when $Similarity\_element > 0.85$ (Equation 1). In searching for elements in proximity, the flow backbone is considered as horizontal or vertical oriented when the standard deviation of E_y or E_x is smaller than $\theta = 0.1$. δ is dynamically changed to the largest distance from backbone to the nearest-K elements ($K = 3$), limited up to 0.2. Two visual groups are considered to be similar if $Similarity\_group > 0$ (Equation 3), i.e., if they have at least one similar element.

To evaluate the performance of our flow construction, 100 infographics were randomly chosen from our inforgraphic collection. Two authors of this article worked together to construct their flow backbones and visual groups manually as ground truth. We then evaluate the performance of the flow construction using the Jaccard Coefficient J between sets of line segments of the constructed and ground-truth backbones. We use the average $Similarity\_group$ between the detected groups and the groups in the ground truth (Equation 3) to evaluate the performance of the group associations. Using our ground truth test set, we get an average $J(backbone, ground\_truth)$ of 0.73, while the average $Similarity\_group$ is 0.61 without considering the position bias in the placement of the amended elements.

Design Pattern Exploration
Using the 13K infographics and their extracted visual information flows, we explored the main VIF design patterns. Inspired by the iterative method of Segel and Heer [34] for exploration of narrative visualization designs, we adopted a semi-automated technique to extract the VIF design patterns from our massive dataset. The basic procedure consists of two parts: (i) the construction of a comprehensive VIF design space and initialization of possible design patterns, and (ii) multiple iterations to construct a categorization of the design patterns.

In the first part, we deployed t-SNE [12] to embed VIFs in a 2D space. For each infographic, its flow backbone and visual group associations are rendered into an image as the *VIF signature* of the infographic (see Figure 1 (left-top)). In the VIF signature, the detected backbones are drawn with thick
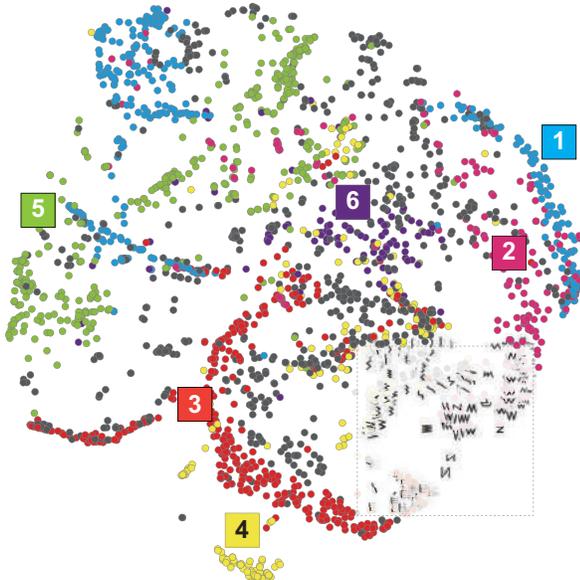
Figure 7. Space of VIF signatures: a down-sampled t-SNE projection with six dominant VIF patterns highlighted (enumerated in Figure 9). A local slicing of VIF signatures is given in the bottom-right.

strokes, while group connections are drawn using thinner lines to distinguish the primary and secondary narrative orders. The signatures of all VIFs have a uniform size of 100×100 pixels. Therefore each infographic is represented by 10K dimensions. To make the analysis practical, we applied Principal Component Analysis (PCA) [22] to extract the top 50 principal components. Taking the 50 vectors as representative dimensions, the VIF signatures were projected to a 2D space using t-SNE (Figure 7).

In the embedding space, similar VIFs are likely to be near each other. Regions in the embedding space that are dense suggest the existence of main VIF design patterns. We applied DBSCAN [14] to spatially cluster VIFs, forming the seeds of potential design patterns. The scatterplot of the identified regions of high density served as the playground to perform the classification. To facilitate and refine the classification, we use an interactive selection tool as a means to examine and define interesting VIFs. The companion video shows a session of VIF signature space exploration.

After this exploratory data analysis, we performed an iterative categorization using the VIF space. We started from seeds with high densities, and spatially expanded them by including neighboring VIFs. This process involved merging similar patterns and re-categorizing divisive patterns. We conducted several iterations until no more significant patterns were found.

VIF TAXONOMY

Following the procedure described above, we identified 12 most notable VIF patterns which are shown in Figure 9. Figure 8 shows their relative prevalence in our dataset.

The VIF patterns are determined according to two main design dimensions: *backbone shape* and *content placement*. The

backbone shape determines the main flow of elements (see Figure 6), while the content placement refers to the way the visual groups are arranged in relation to the backbone.

Figure 8. Relative prevalence of our 12 VIF design patterns.

Backbone Shape. Backbone shape refers to the line that most visual groups in the narrative are aligned with. There are two basic types, circular and linear. Circular information flows, where visual groups are typically aligned around a central object, are further classified by whether they are forming a complete circle or not. VIFs designed in full circles are `clock` and `star`. Patterns designed along an arc are `bowl`, `dome`, `left-wing` and `right-wing` who use a part of the circumference of a circle as the backbone, rotating it to different orientations. Linear information flows arrange visual groups along a line or a curve with explicit order. According to the main orientation, we distinguish between horizontal linear, for which `landscape` is the main example and the vertical linear designs for which `portrait` conveys the VIF whose backbone is a vertical line. `Up-ladder` and `down-ladder` are diagonal designs in-between horizontal and vertical linear, depending on the slope.

Content Placement. Content placement refers to the way visual elements are arranged along the information flow in relation to the backbone shape. In the circular category, the content can be placed inside or outside of the backbone, resulting in `clock` and `star`. In `clock`, the content is placed inside or on top of the backbone, while in `star`, the visual groups are spread outside the center, usually around a central object, with much looser constraints on layout compactness than `clock`. For the horizontal category, the visual groups can be placed on the same side of the flow, or alternatively up and down on different sides. The latter turns out to be the `pulse` pattern. Similarly, in the vertical categories, visual elements can be arranged on one side of the flow which turns out to be the `portrait` pattern, or alternatively on the left and right along the vertical flow, which turn to be the `spiral` pattern.

To better understand the VIF design patterns and its design space for infographics, Figure 10 shows the spatial distribution of three representative graphical data elements, the title, texts for content, and icons for each of the VIF patterns from randomly sampled 2500 infographics.

We can observe the design principle of balancing [4] in these spatial distributions. Elements in the majority of VIF patterns (e.g., `landscape`, `portrait`, etc.) are centered or evenly distributed both vertically and horizontally to achieve *symmetrical balance*, creating a sense of elegance or tight outlook. *Radial balance* is unsurprisingly observed in a circular layout, where the elements are placed circularly around a central point. One observation in the circular layout is that icons are

Figure 9. 12 VIF Design Patterns: six main patterns highlighted in t-SNE plot and six more patterns with variance in backbone and content placement.

usually placed closer to the theme (i.e., the blank area) than texts. We observe *asymmetrical balance* in VIF patterns, such as up-ladder and down-ladder, where elements are usually placed on one side of the flow, off-center, and the title is placed on the opposite side of the main flow to balance the design.

Every infographic has a *theme* or idea that it tries to convey. The theme can be implicitly or explicitly integrated in the infographic. When it is explicitly integrated, it usually appears in the form of a *theme image* - a single image that provides a visual anchor, and usually displays the theme of the info-graphic in a graphical form. With the detected VIF patterns, we approximate the residual region as the potential location of the theme, with a weighted guess according to the specific VIF pattern. For example, the center region of clock is a likely place to hold a theme image. Figure 11 shows the place-ment of the theme image for different VIF patterns (clock, bowl/dome, pulse and landscape from left to right).

Both in the horizontal patterns, such as the landscape or pulse and in the vertical patterns (such as in the portrait or spiral), the theme is usually *implicitly integrated*, not often expressed by a notable thematic image. One exception is in designing the backbone as a theme. For example, see the

rocket figure which implies a 'rising' theme in Figure 2. Still, a theme image can exist in the horizontal patterns, and can be located under or on top of the horizontal flow, while in the vertical patterns, the theme image can be found on the right or left of the vertical flow.

For the circular and semi-circular patterns, a thematic figure is usually *explicitly integrated* at the center of the circular flow, for example, at the center of the star pattern or beneath the dome pattern. The clock pattern can show both implicit and explicit integration. In rounded regions of the clock pattern, the thematic figure is usually centered to emphasize the topic visually. In other cases, the center is left blank to avoid the attention from being attracted by the center.

VIF-EXPLORER
We developed a prototype tool, *VIF-Explorer*, that enables searching infographics according to their VIF patterns. Using the tool, infographics can be searched according to a selected or drawn VIF pattern. In addition, given a specific infographic, other infographics with similar VIF patterns can be retrieved according to their proximity to the selected infographic in the t-SNE projection space. Also, the system supports searching infographics with specified content, such as the number of
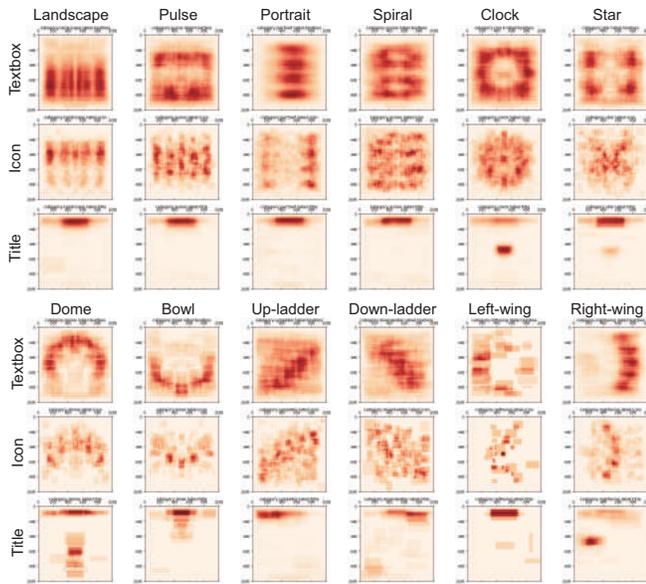
Figure 10. Spatial distribution of title, texts and icons computed from 2500 randomly selected infographics. Note that all infographic are normalized into a uniform size.

Figure 11. Theme Location: the theme image is usually placed in the center of circular patterns, e.g., clock, bowl and dome shown on the left. For linear patterns, while it is less common to have a theme image, if exist, it is placed by shifting the backbone as can be seen in the two images on the right.

visual groups, or whether the infographic includes a specific narrative index. A demonstration of VIF-Explorer is given in the accompanied video. To try the search engine, please visit http://47.103.22.185:8088.

ANALYSIS OF VIF FROM DIVERSE WEBSITES

Our taxonomy of VIF patterns stems from a large design-based infographic dataset as described in the Method section. To examine the coverage and variety of VIF designs that may appear on the Web, we examined our taxonomy with a second set of infographics telling real data stories on various websites. We collected data of general infographics by searching with the keyword 'infographics' on Google Image Engine[4] which retrieves images from various websites. In addition, we also looked at infographics designed for specific domains by searching with keywords 'infographics' together with 'health' in pinterest[5]. For both entries, we discarded images that are clearly not infographics, such as those only with text or only

---

[4]https://images.google.com
[5]https://www.pinterest.com/

graphics, as well as design-based templates without a data story. A total of 965 data-based infographics were collected.

We conducted a multi-pass manual coding of the 965 infographics to identify their VIFs. An interdisciplinary team involving an infographic designer and two authors of this article participated in the coding. In the first pass, each coder went over all the infographics independently and categorized them into 13 categories, the 12 main VIF patterns and one extra category as 'others'. In the second pass, the three coders met and discussed the coding. They consolidated the coding of 12 main VIF patterns by resolving any conflicts. For the infographics categorized into the 'others' category, they examined each one by one to understand why they did not fit any category.

Results of the manual coding revealed that 619 (64.1%) infographics were identified with clear VIF patterns. Among those, spiral and clock are the two most used ones, with (32.3%) and (21.0%), respectively. Portrait is the third with (19.9%), followed by star (8.4%) and landscape (5.8%).
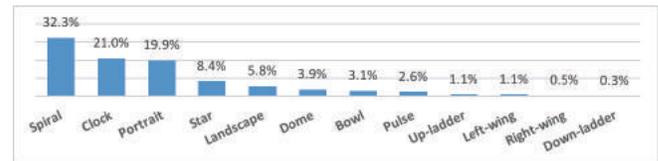

Figure 12. Distribution of 12 VIF patterns identified within 619 Infographics, out of 965 infographics in total (64.1%).

As shown in Figure 12, this distribution of VIF patterns is mostly consistent with our observations of the major VIF patterns found in the main *InfoVIF* analysis (Figure 8). There are dominantly more vertical VIF designs (portrait and spiral) than landscape (landscape and pulse), which might be due to the tendency of documents to be designed and printed in portrait mode. Compared to pure portrait, spiral is more popular in applications, which might be explained by its more interesting visual pattern. In those end-infographics, pulse is surprisingly less used than dome or bowl.

Looking at the 346 infographics in the "others" category, we found that 139 of them (14.4% of the entire corpus) are *compound* infographics, meaning that they are composed of several parts and include two or more of the 12 VIF or graphical patterns (see examples in Figure 13(a)). 20 infographics (2.1% of the entire corpus) are designed with a specific visual information flow, that was not included in our taxonomy (Figure 13(b)). 187 infographics (19.4% of the entire corpus) were found with no distinct order to read (i.e., no VIF). Looking at this group, out of the 187 infographics, 90 are designed in a regular grid layout (see Figure 13(c)), where visual groups are designed with highly uniformed style (e.g., icons in uniform size and texts with the same font) and a regular arrangement. The other 97 infographics are designed without a detectable VIF as can be seen for example in Figure 13(d). Without obvious hints, the readers are free to navigate within the infographics.
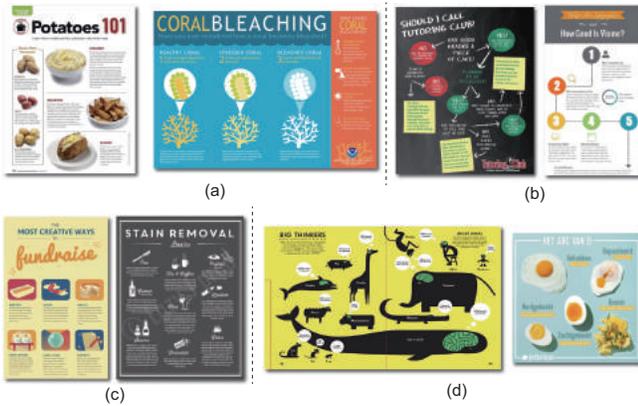
Figure 13. Examples of infographics that do not fall within the 12 patterns of our taxonomy: (a) compound infographics composed of several VIFs; (b) infographics with out of the box information flows; (c) infographics without a VIF, designed in a regular grid layout; (d) infographics flexibly designed without a reading flow.

DISCUSSION

Our taxonomy includes 12 prominent VIF patterns from a broad range of infographics. However, as we saw in the previous section, this taxonomy is not exhaustive, and there are many infographics that cannot be classified within any of these 12 VIF patterns. As Figure 13(a) shows, compound infographics are often designed as dashboards, and consist of multiple parts. The analysis of the 139 compound infographics in the Website dataset, suggests that designers commonly use subtitles, dashed lines, or background colors to visually partition infographics into parts, each of which is designed with a single VIF pattern. In our work, we focus on studying standalone VIFs. One possible research direction in the future is to look into the compound infographics and study how different VIFs are composed together. Besides the structure of visual objects (which we focused on in this work), other visual factors, such as color distinguishing parts in compound infographics, could be further considered in the flow construction.

We derived our taxonomy of major patterns from the space of VIF signatures (Figure 7), according to the major clusters. However, there are some unusual infographics with creative means, for which their visual information flow does not conform with others (Figure 13(b)). Understanding how the information story is narrated in such infographics would be interesting, especially for infographics that guide readers' attention with implicit visual encodings. For example, in some creative infographics, designers might use the artistic decoration to hint the visual information flow, such as a ladder image to indicate the `up-ladder` VIF pattern. How artistic elements help with the visual information flow can be further studied in the future.

During our exploration of VIFs with real-data infographics, the relation between VIF design and the underlying semantics within an infographic was empirically observed. For example, `spiral` or `pulse` is usually used for side-to-side information comparisons. `Star` is a good fit for information which is

around a central idea. While examining these ideas requires further investigation into the semantics of infographics, which is outside the scope of this work, we encourage follow-up research efforts going into this direction.

With the taxonomy of VIFs, another possible follow-up work is to apply these findings to generative tools for creating infographics, for example, by providing VIF patterns as templates for quick infographic generation, or by facilitating the infographic design process with automatic completion of VIF patterns when only a few elements are given.

CONCLUSIONS

In this paper, we introduce a method to explore visual information flows of infographics. Through an analysis of a large scale repository of infographics, we describe a taxonomy of visual information flows. This analysis of unstructured and varied images is made possible thanks to recent advances in neural networks that have now strong competence in understanding images, and in detecting patterns and visual elements. The key is using tools with some preliminary ability to analyze semantics in images. In this work, we leverage these novel abilities to peel off the artistic graphical elements from the data that the infographics carry. It should be stressed that the artistic graphical layer is meant not just to be aesthetic, but also to attract the user's attention and to make the narrative more pronounced. The use of Gestalt principles play a key role allowing the inference of the latent structure of the information flow, using a combination of the proximity, similarity and regularity grouping principles. The Gestalt principles also help to compensate for under (or over) detection of the basic visual elements.

Our technique is not without limitations. There are many outstanding infographics with overly creative means, and their visual information flow does not form a notable pattern with an identified flow signature. Moreover, our VIF design space is analyzed with some user assistance; although minimal, it inevitably injects some subjectivity into the taxonomy. Nevertheless, this work promotes the key idea of mining into the information flow in a broad range of infographics and provides an initial exploration of VIF patterns. We believe that it opens up new avenues and encourages more research efforts in analyzing infographics. We hope that future work will move from analysis to synthesis, offering designers generative tools for transferring styles and patterns from one infographics to another as well as creating novel infographics.

REFERENCES

[1] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Andres Monroy-Hernandez, and Pourang Irani. 2017. Authoring data-driven videos with dataclips. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 501–510.

[2] Benjamin Bach, Zezhong Wang, Matteo Farinella, Dave Murray-Rust, and Nathalie Henry Riche. 2018. Design Patterns for Data Comics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, Article 38, 12 pages.

[3] Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2573–2582.

[4] Jacci Howard Bear. 2019. Find Out If Your Page Layout Is in Balance. https://www.lifewire.com/balance-in-design-1078231. (February 02 2019). Bear2019.

[5] Michelle A Borkin, Azalea A Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. 2013. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2306–2315.

[6] Steven Bradley. 2015. Design Principles: Compositional Flow And Rhythm. https://www.smashingmagazine.com/2015/04/design-principles-compositional-flow-and-rhythm/. (April 29 2015). Bradley2015.

[7] Jean M Brechman. 2010. Narrative "flow": A model of narrative processing and its impact on information processing, knowledge acquisition and persuasion. (2010).

[8] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner. 2017. Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. *IEEE Transactions on Visualization and Computer Graphics* 23, 9 (Sep. 2017), 2151–2164.

[9] Zoya Bylinskii, Sami Alsheikh, Spandan Madan, Adria Recasens, Kimberli Zhong, Hanspeter Pfister, Fredo Durand, and Aude Oliva. 2017a. Understanding infographics through textual and visual tag prediction. *arXiv preprint arXiv:1709.09215* (2017).

[10] Zoya Bylinskii, Nam Wook Kim, Peter O'Donovan, Sami Alsheikh, Spandan Madan, Hanspeter Pfister, Fredo Durand, Bryan Russell, and Aaron Hertzmann. 2017b. Learning visual importance for graphic designs and data visualizations. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, 57–69.

[11] Zhutian Chen, Yun Wang, Qianwen Wang, Yong Wang, and Huamin Qu. 2019. Towards Automated Infographic Design: Deep Learning-based Auto-Extraction of Extensible Timeline. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 917–926.

[12] Laurens Van Der Maaten and Geoffrey E Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9 (2008), 2579–2605.

[13] Agné Desolneux, Lionel Moisan, and Jean-Michel Morel. 2004. Gestalt theory and computer vision. (2004), 71–101.

[14] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, and others. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.

[15] Johanna Fulda, Matthew Brehmel, and Tamara Munzner. 2016. Timelinecurator: Interactive authoring of visual timelines from unstructured text. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 300–309.

[16] Ross Girshick. 2015. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*. 1440–1448.

[17] Steve Haroz, Robert Kosara, and Steven L Franconeri. 2015. Isotype visualization: Working memory, performance, and engagement with pictographs. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 1191–1200.

[18] Lane Harrison, Katharina Reinecke, and Remco Chang. 2015. Infographic aesthetics: Designing for the first impression. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1187–1190.

[19] Jeffrey Heer and George Robertson. 2007. Animated transitions in statistical data graphics. *IEEE transactions on visualization and computer graphics* 13, 6 (2007), 1240–1247.

[20] Weiyin Hong, James YL Thong, and Kar Yan Tam. 2004. Does animation attract online users' attention? The effects of flash on information search performance and perceptions. *Information Systems Research* 15, 1 (2004), 60–86.

[21] Jessica Hullman, Steven Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. 2013. A Deeper Understanding of Sequence in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2406–2415.

[22] Ian Jolliffe. 2011. Principal component analysis. In *International encyclopedia of statistical science*. Springer, 1094–1096.

[23] Nam Wook Kim, Zoya Bylinskii, Michelle A Borkin, Krzysztof Z Gajos, Aude Oliva, Fredo Durand, and Hanspeter Pfister. 2017. BubbleView: an interface for crowdsourcing image importance maps and tracking visual attention. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 5 (2017), 36.

[24] Nam Wook Kim, Eston Schweickart, Zhicheng Liu, Mira Dontcheva, Wilmot Li, Jovan Popovic, and Hanspeter Pfister. 2016. Data-driven guides: Supporting expressive design for information graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 491–500.

[25] Ranjitha Kumar, Arvind Satyanarayan, Cesar Torres, Maxine Lim, Salman Ahmad, Scott R Klemmer, and Jerry O Talton. 2013. Webzeitgeist: design mining the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3083–3092.

[26] Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. 2013. SketchStory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2416–2425.

[27] Thomas F Liu, Mark Craft, Jason Situ, Ersin Yumer, Radomir Mech, and Ranjitha Kumar. 2018a. Learning Design Semantics for Mobile Apps. In *The 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, 569–579.

[28] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. 2016. Ssd: Single shot multibox detector. In *European conference on computer vision*. Springer, 21–37.

[29] Zhicheng Liu, John Thompson, Alan Wilson, Mira Dontcheva, James Delorey, Sam Grigg, Bernard Kerr, and John Stasko. 2018b. Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Expressive Visualization Authoring. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, Article 123, 13 pages.

[30] Sean McKenna, N Henry Riche, Bongshin Lee, Jeremy Boy, and M Meyer. 2017. Visual narrative flow: Exploring factors shaping data visualization story reading experiences. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 377–387.

[31] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779–788.

[32] Babak Saleh, Mira Dontcheva, Aaron Hertzmann, and Zhicheng Liu. 2015. Learning Style Similarity for Searching Infographics. In *Proceedings of the 41st Graphics Interface Conference (GI '15)*. 59–64.

[33] Arvind Satyanarayan and Jeffrey Heer. 2014. Authoring Narrative Visualizations with Ellipsis. *Computer Graphics Forum* 33 (2014), 361–370.

[34] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.

[35] Teo Siang. 2019. The Building Blocks of Visual Design. https://www.interaction-design.org/literature/article/the-building-blocks-of-visual-design. (March 2019).

[36] Martin van den Berg, Michael Kubovy, and James A Schirillo. 2011. Grouping by regularity and the perception of illumination. *Vision research* 51, 12 (2011), 1360–1371.

[37] Yun Wang, Haidong Zhang, He Huang, Xi Chen, Qiufeng Yin, Zhitao Hou, Dongmei Zhang, Qiong Luo, and Huamin Qu. 2018. InfoNice: Easy creation of information graphics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 335.

[38] Nanxuan Zhao, Ying Cao, and Rynson WH Lau. 2018. What characterizes personalities of graphic designs? *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 116.