Quality Metrics for Information Visualization

M. Behrisch[1], M. Blumenschein[2], N. W. Kim[1], L. Shao[3], M. El-Assady[2], J. Fuchs[2], D. Seebacher[2], A. Diehl[2]

U. Brandes[4], H. Pfister[1], T. Schreck[3], and D. Weiskopf[5], D. A. Keim[2]

[1] Harvard University Cambridge, USA
[2] University of Konstanz, Germany
[3] Graz University of Technology, Austria
[4] ETH Zürich, Switzerland
[5] University of Stuttgart, Germany

Abstract

The visualization community has developed to date many intuitions and understandings of how to judge the quality *of views in visualizing data. The computation of a visualization's quality and usefulness ranges from measuring clutter and overlap, up to the existence and perception of specific (visual) patterns. This survey attempts to report, categorize and unify the diverse understandings and aims to establish a common vocabulary that will enable a wide audience to understand their differences and subtleties. For this purpose, we present a commonly applicable quality metric formalization that should detail and relate all constituting parts of a quality metric. We organize our corpus of reviewed research papers along the data types established in the information visualization community: multi- and high-dimensional, relational, sequential, geospatial and text data. For each data type, we select the visualization subdomains in which quality metrics are an active research field and report their findings, reason on the underlying concepts, describe goals and outline the constraints and requirements. One central goal of this survey is to provide guidance on future research opportunities for the field and outline how different visualization communities could benefit from each other by applying or transferring knowledge to their respective subdomain. Additionally, we aim to motivate the visualization community to compare computed measures to the perception of humans.*

1. Introduction

The idea of measuring the quality of a visualization is as old as the information visualization community itself. Early work in the field can be traced back to the work of Bertin [Ber81], although the notion and importance of quality were developed far earlier in cartography. Undoubtedly, Tufte was the first research pioneer formalizing the quality metric idea to a simple, thus understandable quality metric: the *data-to-ink ratio* [TGM83]; a metric to convey the core principles of an effective and efficient, crisp design.

Generally, effective and efficient visualizations follow a simple mantra: They show the most information in the simplest possible form. However, the current data to be visualized puts more and more challenges on visualization designers: high-dimensional s-paces, complex relationships, or the sheer amount of data to be visualized demand a careful choice of the visual variables for a *faithful* representation of the underlying dataset.

Following the accepted *information visualization pipeline* of Card et al. [CMS99] –as one possible example– a visualization designer will inevitably be confronted with the dilemma of choosing from a multitude of data processing possibilities and an even greater choice of potential visualization options. To give a practical example: If a user wishes to visualize a 20-dimensional dataset, not only data-specific questions, such as normalization and outlier removal, play a critical role, but also which data characteristic should be highlighted first. In case that a visualization designer decides for a Scatter Plot, which fixes most of the choices of the visual variables, $n \times (n-1)/2$ potentially meaningful dimension combinations can be depicted. Each of these 190 views needs to be evaluated independently for its usefulness by analyzing its effectiveness concerning other visual encodings, such as color mapping, visual marks, and axis ranges.

In the general case, the number of visual mappings for an arbitrary data type grows exponentially with the number of mapping options, thus making information visualization design to a *trial-and-error process*. More importantly, however, is that only those visualizations can be considered effective that support the building of mental models for the underlying dataset [Nor06]. Hence, the essence of effectiveness resides in the identification of *interpretable visual patterns* that contribute to the overarching analysis goal.
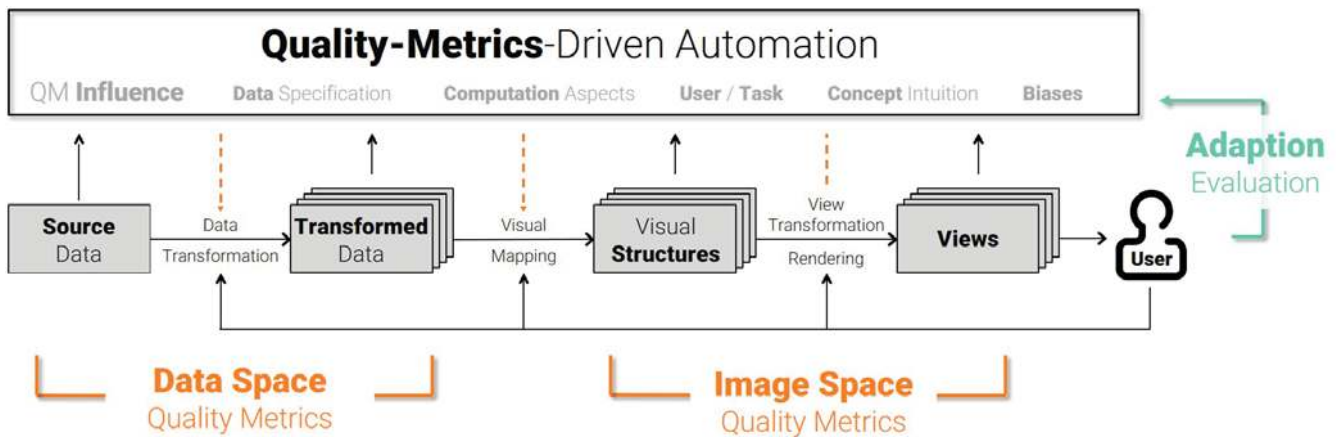
Figure 1: *Quality Metrics-driven Visual Analytics pipeline. The pipeline adds an additional layer named Quality Metrics-driven Automation on top of the traditional information visualization pipeline [CMS99]. The layer could obtain information about the several stages of the pipeline (the boxes) and influences the processes of the pipeline through the quality metrics it calculates. The user is always in control. Image and text adapted from [BTK11].*

The research field of Quality Metrics (QMs) has devoted its efforts to develop *quantitative measures* for detecting visualizations that contain one or multiple interpretable visual patterns. Applied to exploration and navigation contexts, quality metrics can help to guide the user to views of interest or can help to mitigate the cognitive overload by filtering cluttered or uninteresting views. In general, quality metrics stand as an umbrella term for quantifying the (visual) quality and such the effectiveness and interestingness of a visualization. These approaches find broad applications in the visualization of high-dimensional, relational, or geospatial data. Over the last 30 years, a myriad of approaches, techniques, and concepts have been developed to help the user find a suitable data transformation and visual mapping by iterating and evaluating every possible visualization design combination.

Our motivation for this report is two-fold. First, we recognize that by now the most recent quality metrics surveys date back several years [BTK11, ED07]. In the meantime, the field was undergoing an important development from quality metrics that heuristically quantify the amount of clutter toward a pattern- and analysis task-driven exploration. Therefore, we aim to provide an update by adding more recent publications to the body of work presented in these earlier surveys. Second, we noticed that, although a wide range of approaches was presented under the headline of quality metrics, only little effort has been devoted to describing the methodological and conceptual background of these approaches. Consequently, this work aims to bring depth into the discussion, by consistently enumerating, describing, and relating the underlying concepts with the same vocabulary. As the third motivation point, we claim that most approaches have not yet been evaluated for their perceptual relationships. However, novel and innovative evaluation approaches, such as crowdsourcing and hardware developments (eye trackers in a sub 100$ range) are opening new potentials for this research field.

In summary, the contribution of this paper is to give a comprehensive overview of existing quality metrics for different information visualizations techniques, particularly Scatter Plots (6.1) and Scatter Plot Matrices (6.2), Parallel Coordinates (6.3), Pixel-based Techniques (6.4), Radial Visualizations (6.5), Glyphs (6.6), Node-Link Diagrams (7.1), Matrix Representations (7.2), TreeMaps (7.3), Line Charts (9.1) and Stacked Graphs (9.2), Typographic Visualizations (10.1) and Tag Clouds (10.2), and Geo-Spatial Data Visualizations (8). Our selection is targeted towards fields in which QMs are in focus of the research, but we also outline a potential usefulness of QMs for other visualization techniques. As a guiding theme, we not only concentrate on a pure enumeration of quality metrics techniques but focus more on a detailed description of the underlying *concepts* and *models* and their variety of different implementation possibilities. Furthermore, we survey how QMs are evaluated and whether results are compared to the human perceptiveness.

2. Background and Conceptualization

This section introduces definitions and concepts that we rely upon to describe quality metrics approaches. We discuss common concepts and methodologies across different visualization domains. As one of the core motivations of this survey, we plan to unify the vocabulary and understanding of quality metrics. To achieve this bold goal we gradually increase the level of formalism in the following section. To ease the readability we decided to begin with a purely informal description of our *quality metric vision*. Then we present our attempt to formalize the problem and describe thoroughly constituents and facets influencing the understanding of quality metrics.

2.1. Quality Metric Vision

The grand and sketchy vision behind the visual quality metrics research is the following: Imagine a visual analysis would be based on a black-box that that is fed with your current analysis task(s), user preferences, and the dataset at hand. This black-box would "auto-magically" derive a recommendation of the best possible visualization type and visualization instantiation; would derive the most effective visual variable settings (e.g., color map, shape, texture) and all necessary data preprocessing steps depending on multitude

of soft and hard influencing factors; and would finally present the most interesting view on the data that reveals most information.

However, while this vision sounds overarchingly promising, parts of the questions can already be tackled with current technologies: More or less sophisticated "Show Me" buttons (e.g., [MHS07]) decide for the user which visualization is appropriate based on data types. Other approaches even add considerations about the underlying data distribution into their recommendations of a visualization type and visual mapping [WMA*16, WQM*17].

Other approaches start from the constraint that the visualization type is fixed, e.g., Scatter Plots for projections of high-dimensional data and tackle the question which views can be discarded due to the high overlap or *visual clutter* [BS04, TBB*10]. Again other approaches, such as the so-called *-gnostics [WAG05, SSK06, DK10, LKZ*15, BBH*17], focus on the quantification of visual patterns for their specific visualization type, following the core idea of promoting only views containing interpretable visual patterns and thus helping build mental models about the dataset and task relationships.

But, while we are seeing more and more advanced research for supporting the user in the exploration process, the current research is struggling with the definition, categorization, and labeling of the current exploration task in place. Partially this problem arises because exploration tasks are not necessarily separable in terms of their temporal characteristics and oftentimes even nested in nature. To make matters worse, most users do not follow a structured exploration path but conduct several exploration tasks in parallel with a more or less prominent specificity. While basic research has been presented in this field, such as various task taxonomies with different levels-of-details [BM13, KAF*08, LPP*06, Shn96], only a few works focused on automatically quantifying the current exploration task at hand. Sacha et al. [SSZ*17] list a range of works following this research stream. Additionally, quality measures should approximate the users' perception and cognition. Yet, only a few approaches have been evaluated with user studies and only a few evaluations compare the usefulness of multiple different metrics.

The aforementioned consideration sketches outline a far-reaching and extensive research field with multi-faceted foci and research potential for at least the next decade. Consequently, we will not be able to report on all developments. Rather, we decided to put emphasis on what we denote as *Mid-level Perceptual Quality Metrics*. This emerging field focuses on perceptually-inspired quality metrics that try to mimic parts of the human perception/cognition in order to ease the exploration process. These approaches not only reduce the cognitive overload by separating the "wheat from the chaff", i.e., by removing noise, but also facilitate building task-related mental models by mimicking the humans' ability to recognize and differentiate between visual patterns.

Exemplified Mid-level Perceptual Quality Metric-Driven Exploration Workflow

In an exemplified usage scenario for a *Mid-level Perceptual Quality Metrics*-driven exploration, a user would have to analyze a large set of matrix plot images for the visual patterns they contain. These scenarios occur regularly, e.g., in the medical data domain, where brain activity maps are generated in millisecond time intervals and —for example— deviations from the baseline brain activity maps are to be retrieved. Another application scenario is the performance analysis of massively parallel computing systems. One application run on

a High-Performance Computing Cluster can produce many time-dependent performance measures in (sub-)millisecond intervals for many clusters, containing many computing nodes. Understanding these large multiplicative exploration spaces becomes extremely time-consuming, if not impossible. In a quality metric-driven exploration, the user would try to reduce the cognitive overload by computationally assessing the interestingness/effectiveness of each view. This can be facilitated in two ways: clutter reduction approaches will discard all matrix views with a low signal-to-noise ratio, while pattern-driven exploration approaches would show the user the distribution of visual (anti-)patterns in the view space. If the user's task is explicit and well-specified, such as finding evidence of a data partitioning/grouping, pattern-driven navigation concepts can outperform clutter-reduction approaches, because in our case the user would just have to search for block-diagonal matrix plots.

2.2. Definitions

In the remainder of this paper, we use the following definitions. Formally, measuring the quality of a visualization V consists in computing one visualization definition $\phi \in \Phi$ from a universe Φ of potential instantiations that maximizes or minimizes a specified *quality criterion* $q(D, U, T)$, such that:
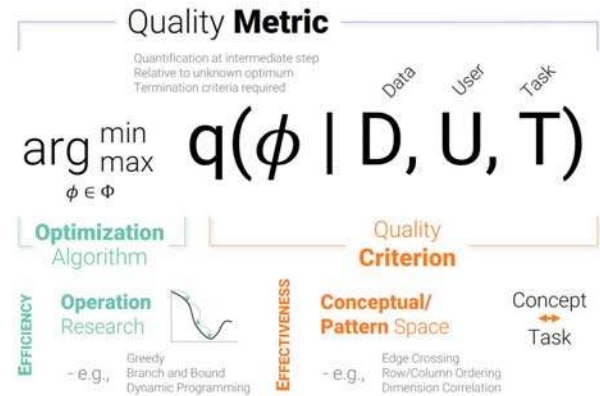

Figure 2: *Quality Metrics (QM) formalization. QMs are composed of an algorithmic part and a quality criterion. A potential multi-objective optimization algorithm tries to find efficiently a valid visualization configuration (ϕ) that optimizes the designed quality criterion $q(\phi \,|\, D, U, T)$. The quality criterion tries to heuristically capture how an effective visualization instance might look like. This intuition is bound and influenced by the task T at hand (defines the to-be-expected visual appearance), the dataset characteristics D (defines if a visual pattern is producible), and the user preferences U. Consequently, a QM arg min/max $q(\phi \,|\, D, U, T)$ determines a perceptually preferable visualization configuration ϕ for a given quality criterion $q(\dots)$ given the influencing factors D, U, T.*

where D denotes the dataset, U the user and his/her current capabilities and T the user's current analysis task at hand.

To illustrate our formalism let us imagine the following scenario: We describe our user U as a statistically knowledgeable person with average attention potential whose task T is to understand data/dimension (dis-)similarities in a high-dimensional dataset D. Our
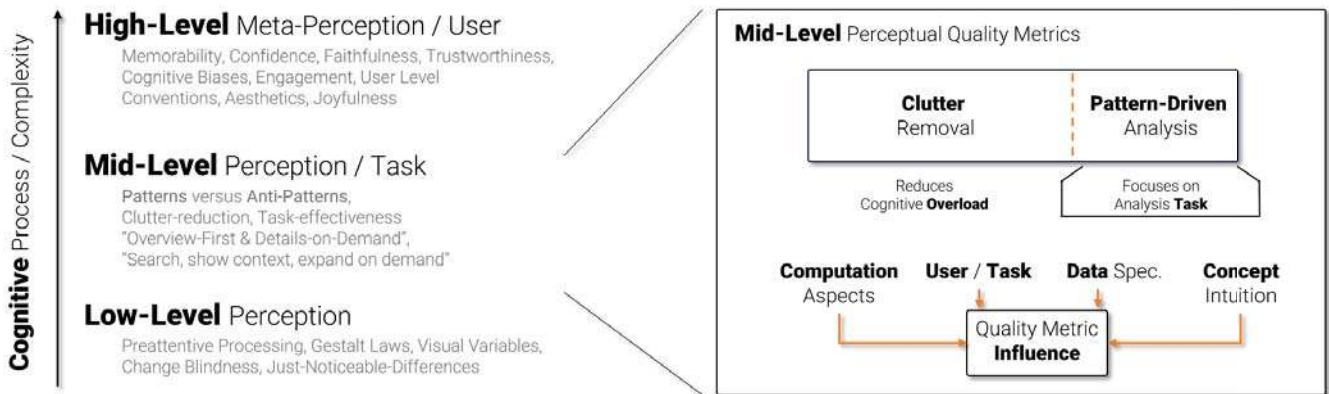
Figure 3: *A wide range of quality metric understandings exist in the literature. The left side shows a broad categorization of the field sorted according to the cognitive complexity these approaches try to reflect. The right side shows that the task focus of* Mid-level Perceptual Quality Metrics *comprises different granularity levels: (1) Overview: distinguish between noise/clutter and any kind of pattern, (2) quantify the quality of a visualization based on a specific pattern (depends on the analysis task).*

Quality Metric-driven recommendation system could decide that a Scatter Plot display is a suitable choice to show (dis-)similarities for this kind of user. The quality criterion $q(\dots)$ could then compute the sum of pairwise distances over all displayed points in D with respect to a chosen distance function while taking the data specifics into consideration (i.e., needs outlier cleaning). The Equation in Figure 2 would find for a specific task $t_1 \in T$, a $\phi_1 \in \Phi$ that minimizes this sum relating to a locally dense Scatter Plot or could find $\phi_2 \in \Phi$ that maximizes the sum to find globally cluttered plots for another task $t_2 \in T$.

Mid-Level Perceptual Quality Metrics

The area of Mid-level Perceptual Quality Metrics leaves out all considerations about the user U; assessing his/her skill set or cognitive/physiological capabilities and does not (yet) deal with an explicit formulation of tasks T during the exploration process. The field of Mid-level Perceptual Quality Metrics is rather concerned with presenting heuristics and algorithms to statistically quantify the extent of an anti-pattern –e.g., how a cluttered view looks like– or which specific visual pattern is apparent –e.g., locally dense Scatter Plots can be used to reason about data similarity.

In the following, we will outline the components contributing the definition of a Mid-level Perceptual Quality Metric.

(i) A **Quality Metric (QM)** combines an optimization algorithm and quality criterion with the overarching goal to mimic parts of the human perception. QMs are developed with a specific goal in mind, such as finding clutter-free visualizations or visualizations with a specific interpretable visual pattern.

(ii) **Visualization Definition** ϕ is an instantiation of the parameter space Φ defining the appearance of a specific visualization type. Following the information visualization of Card et al. [CMS99], as depicted in Figure 1, we will have to distinguish between *data*-dependent and *visualization*-dependent parameters. For a Scatter Plot, ϕ would define the necessary data transformations, such as which outliers will distract the view "too much" and the view-space parameters describing the visual appearance of data item (e.g., shape, color, texture,

position) and the corresponding axis definition and appearance (e.g., offset, normalization type, aspect ratio).

(iii) **Quality Criterion** $q(\dots)$ is an (heuristic) algorithm or function for quantifying the effectiveness of one visualization instantiation/view. In other words, a quality criterion evaluates heuristically whether or not a view follows established perceptual guidelines. In the most cases, the goal is to quantify the visual appearance of (anti-)patterns. We consider visual patterns as the target elements of the exploration process, while visual anti-patterns, such as noise, will distract the user without adding to his/her understanding about the dataset and task at hand.

(iv) **Optimization Algorithm** makes use of a quality criterion and -concept to derive, e.g., a ranked or filtered list of visualization instantiations (or views). To achieve this goal an optimizer takes a quality criterion and improves the measure over the visualization method parameters ϕ. Most prominently, filtering concepts are applied to discard cluttered views, while pattern-exploration systems categorize views in terms of the visual patterns they contain.

Note that *metric* has a precise meaning in mathematics, but is used more loosely in the present context. The characteristics of a metric, i.e., non-negativity, identity of indiscernibles, symmetry, and the triangle inequality, need not necessarily hold in all cases. As an example, many QM approaches are based on non-deterministic computations to retrieve (good) local optima in the visualization parameter space. Hence, the term quality metric should be rather understood as an artifact that developed over time from a mathematical understanding toward a more vague and indistinguishable field of more or less mathematically backed up research approaches.

2.3. Common Calculation Approaches

In our literature review, we identified three different concepts to compute quality metrics: a primarily *image space* dependent computation, a purely *data space* dependent computation, and *hybrid approaches* that efficiently combine both concepts. Moreover, we found that QMs are either used implicitly during the construction of

visualizations or as a separate evaluation component complementing the construction and use of visualizations.

Image Space QMs assess the quality of a visualization solely based on the rendered image. Often, sophisticated feature descriptors are extracted from the image and used to measure clutter or perceivable patterns. For example, Tatu et al. [TAE*09] encode the visual quality of Parallel Coordinate Plots by means of a Hough Space feature descriptor. With this approach, it is possible to distinguish visually noisy and strongly clustered axis combinations. In a quality metric driven analysis, we aim to mimic the perception of a human to identify patterns. The main advantage of an image-based quality assessment is therefore that we use the same visual information (i.e., image) that is also assessed by humans in an evaluation setting.

Data Space QMs measure the quality of a visualization before the rendering process starts. The approaches are based either on raw or transformed input data, or estimate how visual structures will most probably look like. As an example, Johansson and Johansson [JJ09] propose an interactive approach to weight multiple data spaces based quality metrics to reorder Parallel Coordinate Plots. Their metrics comprise a user-defined weighting of correlation dimensions (by a Pearson correlation coefficient), outlier analysis (by a grid and density based approach), and cluster detection (by applying a subspace clustering algorithm). The main advantage of data-based QMs is that many measures (such as cluster algorithms) exist and can be computed usually quite efficiently.

Hybrid QMs combine the advantages of image and data space approaches. For example, Bertini and Santucci [BS04] determine a good sampling rate in Scatter Plots by comparing the visible data density in image space with the relative data density in data space. The number of visible points at one specific location in the visualization is either 0 or 1 in the image space, while the data space can also count more than one points at one location. Combining these measures support most useful sampling strategies.

Implicit / Explicit QMs Many approaches make use of implicit quality criteria *as part of* an optimization problem. Typically, these approaches do not explicitly externalize numeric scores for the quality of a visualization, but decide during the view construction which representations is more useful.

To bring a practical example, the ordering of rows and columns in a matrix visualization lets visual and interpretable patterns either emerge or disappear. Thus, matrix reordering incorporates an implicit notion or understanding of the expected patterns. Another classical example in the context of dimension reduction is presented by Wang et al. in [WFC*18]. For labeled datasets, typically depicted by color-coded Scatter Plots, they start with a (pseudo-)random placement of items in 2D. This placement is incrementally improved wrt. one or multiple visual class separation QMs by choosing the one perturbation of the current solution that improves the QMs. Integrated into a simulated annealing optimization, this approach helps to traverse the exploration space and find a locally optimal solution for the chosen class separation QMs.

An explicit quality criterion for matrix patterns would quantify *to which extent* specific visual patterns are present in the current matrix view described by its reordering algorithm. But, explicit QMs can also be used to choose *between* various visualization types and configurations. For example, in "Line Graph or Scatter
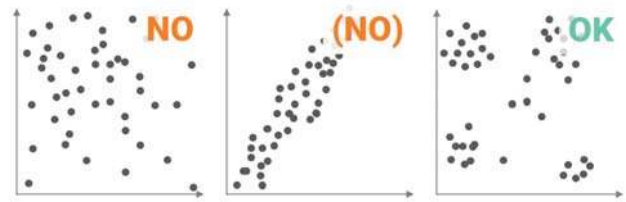

Figure 4: *Example of a* task-specific quality metric *for Scatter Plots. Task: finding data groupings or clusters.*

Plot? Automatic Selection of Methods for Visualizing Trends in Time Series" [WHZ*18] the authors quantify the visual consistency between the data set's trend curve and the trend described by a scatter plot or a line graph. Based on the numeric comparison of both QM scores, the better visual approximation is chosen.

2.4. Analysis Scenarios Supported by Quality Metrics

We can distinguish between QMs designed for *clutter reduction* and *pattern-driven analysis*, as depicted in Figure 3. Clutter reduction techniques reveal the contained set of visual structures by "only" filtering out noisy views. Therefore, they are most useful to obtain an overview of large and unknown datasets, as they keep all views with potentially interesting visual patterns. Hence, these QMs mitigate the cognitive overload problem. However, users typically have specific exploration or analysis foci in mind to understand the data structure and topology. Searching for visual patterns with particular properties is significantly more challenging and requires a quantification and distinguishing of visual structures. But, perceptually-inspired QMs have the benefit to support the user *directly* by contributing to their mental model and understanding of the data.

Overview of Analysis Tasks. Quality metrics identifying a particular pattern are typically related to one or more analysis tasks. We refer to these metrics as *task-specific quality metrics*. For all QM that we report in this paper, we try to elaborate on the (potentially) underlying task(s). We do not stick to any of the established task taxonomies, since they are too specific compared to the analysis tasks supported by QMs. In contrast, we present a high-level overview of exploration tasks supported by the majority of metrics:

 (i) **Clutter reduction.** Users are interested in filtering out noisy views without a specific visual pattern in mind. This task is a typical used to get an overview of unknown datasets.

 (ii) **Preservation task.** QMs for preservation tasks identify views that preserve the original data properties in the mapping process. The preserved aspects can be, e.g., individual data points, topological structures, or distance in map distortions.

(iii) **Search for data groups and partitions (clusters).** QMs aim to identify views in which a (useful) partition and/or dense groups of data records are visible.

(iv) **Search for outliers.** The goal is to identify views that highlight data points differing from the majority of other points. Hereby, the notion of "difference" depends on the application.

 (v) **Search for dimension relations.** Views supporting this task depict combinations of dimensions showing relationships between the data points (e.g., correlations).
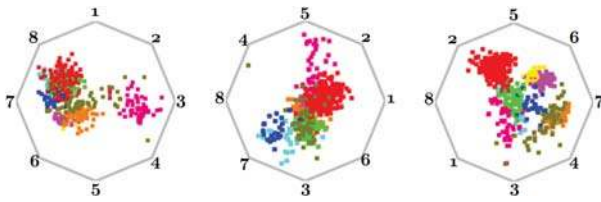
Figure 5: *Different dimension orderings in RadViz preserve and emphasize, respectively mask, given groupings in high-dimensional data. Groupings become increasingly visible from left to right. Figure adapted from [AEL*10].*

Figure 6: *Average income of persons in the US represented on a Map. A linear mapping between income and color (left) and logarithmic mapping (right) is used to emphasize data characteristics. Figure adapted from [SSK06].*

(vi) **Data and visualization specific tasks.** For one data type, different visualization techniques exist; each with (dis-) advantages to reveal essential aspects. Some analysis tasks are specific to data or visualization types (e.g., readability of typographic visualizations) and cannot be generalized.

One example of a task-specific QM is shown in Figure 4. Imagine an analysis task in which users need to find data groupings (clusters) in Scatter Plots: While the first Scatter Plot contains only noise, the last plot reveals several clusters, detected by a quality metric. Although the second plot also shows an interpretable pattern (correlation of the data), it is not relevant to the current task. A task-specific quality metric needs to classify the plot as non-interesting due to the non-relevant visible pattern. One existing quality metric that can distinguish between a variety of patterns in Scatter Plots is Scagnostics [WAG05]. It captures the presence of the following nine visual features: outlying, skewed, clumpy, convex, skinny, striated, stringy, straight, and monotonic. In the example we would search for Scatter Plots with low monotonic and high clumpy features.

In another example, an analyst wants to measure how much information is preserved by projecting a high-dimensional dataset with class labels into a 2D representation. The analyst decides to use the RadViz technique and represent color with the class information. As shown by Figure 5, a task-dependent quality metric can help to optimize the ordering of dimensions such that the provided classes are well separated. A quality metric that facilitates this concept is presented by Albuquerque et al. [AEL*10]. Their approach is to measure the density of all classes in every 2D representation.

In a third example, a quality metric can support the selection of an appropriate colormap and/or normalization strategy to identify data groupings and outliers in pixel-oriented techniques. Figure 6 visualizes the counties of the US. Color shows the average income per person. In the left map, a linear mapping between the average income and color is applied, while a logarithmic mapping is used for the right map. Although the logarithmic mapping distorts the real income values, it helps to identify patterns (e.g., high income in the areas of the east and west coast, a rather low income on a horizontal axis in the middle of the US). Eisemann et al. [EAM11] presents a data driven method to decide for an appropriate color mapping to reveal patterns as required above.

During our literature review, we recognized that a some tasks are well-supported by QMs, while others are not. We discuss well-adopted tasks in their respective visualization section and point to open research gaps in these section's open research discussion.
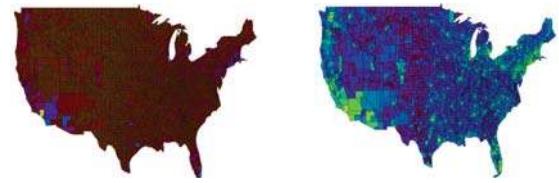
3. Related Concepts

As mentioned earlier this work surveys the recent advances and state-of-the-art for mid-level perceptual QMs. However, this subfield is embedded into an overall quality metric landscape, depicted in Figure 3. For the sake of completeness and delineation, we will enumerate the main concepts and relationships in this section.

The topic of quality metrics is not described in technical terms, but rather incorporates a wide range of understandings. Since one of the core contributions of this paper is to establish a common vocabulary, we are categorizing QM related concepts along the axis of cognitive complexity.

3.1. Low-Level Perceptual Quality Metrics

Low-level perceptual quality metrics leverage the low-level processing of visual stimuli in human perception system such as preattentive processing [War13, HE12]. They are concerned with how basic visual encoding variables, such as position, length, area, shape, and color, and the interaction of the variables (e.g., integrable or separable) influence the efficiency of low-level perceptual tasks such as visual search, change detection, and magnitude estimation.

A great deal of prior visualization research has been devoted to understanding the effectiveness of different visual variables for encoding quantitative and qualitative data. For example, Cleveland and McGill [CM84] ran a series of graphical perception experiments to measure accuracy in comparing values and to derive the rankings of encoding variables. Similar experimental methods have been frequently employed to compare different chart types as well. The results of such experiments have also played a vital role in the automatic construction of visualizations [Mac86, MHS07].

3.2. High-Level Perceptual Quality Metrics

High-level perceptual quality metrics refer to cognitive measures such as memorability, aesthetics, and engagement [SES16]. While they are often considered as subjective dimensions of visualization design, recent studies attempt to quantify these measures based on experiments with human subjects. For example, Borkin et al. [BBK*16] showed that visualization memorability is consistent across people, suggesting that some visualizations are more memorable than others independent of subjects' context and biases.

Various factors can contribute to high-level perceptual quality metrics such as visual density and human recognizable objects for memorability [BBK*16], colorfulness and visual complexity for aesthetics [HRC15], and amount of interactions for engagement [SES16]. While mid-level and low-level perceptual quality metrics

tend to focus on optimizing performance measures for data exploration and analysis tasks, high-level perceptual quality metrics put more emphasis on enhancing the communication aspect of visualization (e.g. whether a visualization can attract the attention of an audience and get the message across).

3.3. Design Recommendations

We consistently recognize two "end products" for quality metric design in all quality metric subfields:

(i) In some visualization subfields, QM results are communicated via **Design Recommendations**; textual guidelines and arguments summarizing the findings about visualization design mostly derived from user-studies. Design recommendations have the great advantage that they represent reproducible evaluations of how a human perceives a view. They can summarize complex perceptual circumstances. Their biggest disadvantage is that these textual guidelines are often derived from simplified task- and context settings that often cannot be generalized to real-world environments and problem settings.

(ii) In other visualization subfields, purely **Heuristic Approaches** prevail. These algorithms model some form of understanding of how a visualization should look like in order to be effective/useful. The biggest advantage of heuristics is their reproducibility, thus allowing user- and context-independent, quantitative visualization comparisons. Their biggest negative point is that visualizations are often judged for their perceptual quality with quantitative scores that have never been proven to correspond to the humans' judgment.

We claim that both approaches are valid but should eventually be backed up with the other approach. Heuristics should be evaluated for their perceptual aspects and proven to be perceptual; design recommendations should be developed into quantifiable heuristics to allow for fair and quantitative evaluation schemes.

4. Related Work

Quality Metrics have been developed for different information visualization techniques. From a historical perspective, we are inspired by a range of survey works with a more or less specific notion of quality metrics. For example, Brath [Bra97] described several image space quality metrics, such as occlusion percentage or percentage of identifiable points, to assess the quality of business visualizations. Miller et al. [MHNW97] expressed the need for new metrics to compare visualizations. Similarly, Diaz et al. [DPS02] advocated the use of implicit and explicit quality metrics for assessing the quality of vertex ordering approaches. In this context, the term of aesthetics is used as same as it is traditionally used in the graph drawing community and refers to a set of measures to reduce the cognitive load for graph exploration tasks [DBETT94, WPCM02].

A first survey focusing primarily on quality metrics for Scatter Plots and Parallel Coordinates was presented by Bertini et al. [BTK11]. Similar to our approach, their survey presents a systematic analysis focusing on the guiding questions: (1) What was measured? (2) Where was it measured (data/image space)? (3) What is the purpose of the QM? And, (4) does the QM allow to be interactively adapted? In total, 20 papers are surveyed in this work.

The evaluation of quality metrics has gained increasing importance in the recent years. For example, Lehmann et al. [LHT15] and

Pandey et al. [PKF*16] study independently the questions about the connection of human perception and (heuristic) quality metrics and present both crowdsourcing studies to prove evidence that this connection exists. Sedlmair and Aupetit [SA15] even present a data-driven framework for quality measure evaluation. Their approach tries to mitigate the impact of (relative) human judgments by relying entirely on ground-truth data. However, this in turn also indirectly implies some sort of user involvement.

An information theoretic approach for assessing the effectiveness of information visualization has been mainly pursued by Chen et al [CJ10]. They built on the initial work by Yang-Peláez et al [YPF00] and proposed a number of entropy-based measures, including visual-mapping ratio, information loss ratio, and display space utilization; these measures are akin to the data-ink ratio [TGM83]. Chen et al also discussed visual multiplexing [CWB*14] in relation to the information theoretic measures. They describe various mechanisms for overlaying multivariate data and discuss how to overcome perceptual difficulties such as occlusion and cluttering that arise from the interference among spatially overlapping visual channels. We consider that these measures concern low-level quality metrics and thus are not discussed in this paper.

Saliency-based measures for evaluating the visualization quality have gained recent interest. They assess how well visually salient regions in a visualization can help users accomplish their goals and tasks. For instance, Chen and Jänicke [JC10] proposed a method for computing a saliency-based metric to measure the mismatches between visual salience and data characteristics (e.g., features detected by algorithms). Matzel at el [MHD*18] recently developed a saliency model to predict where people would look for a given visualization. Unlike models designed for images of natural scenes, their model attempts to incorporate top-down visual features (e.g., texts) that are crucial for visualization tasks. Tailoring the models for different visual analysis tasks is largely unexplored, however [PWV*18]. We believe that saliency-based measures touch on both low-level and high-level quality dimensions and thus not addressed in this paper.

Although the field of quality metrics for color mapping can be safely categorized into low-level perceptual quality metrics research and is thus not in the focus of this survey, we decided to stress some shared argumentation paths by selectively summarizing some more recent works. Quality metrics for color mapping have been investigated amongst others in the work of Bernard et al. [BSM*15], Mittelstädt et al. [MJSK15, MK15], or recently by Gramazio [GLS17]. Szafir and Gleicher [SG16] argue for choosing colors based on a given context rather than in isolation. They identified three categories of design constraints and make design recommendations for effective color choices based on aesthetic constraints, perceptual constraints, and functional constraints. Eisemann et al. [EAM11] present an orthogonal approach. Based on a range of data analysis and transformation steps, a user-independent, data-driven color mapping approach is postulated.

While many approaches are targeted toward clutter removal [ED07], only very few are targeted toward describing the perceived appearance with respect to *visual patterns*. Our survey aims at describing quality metric approaches in a unified manner to better understand their differences and subtleties.
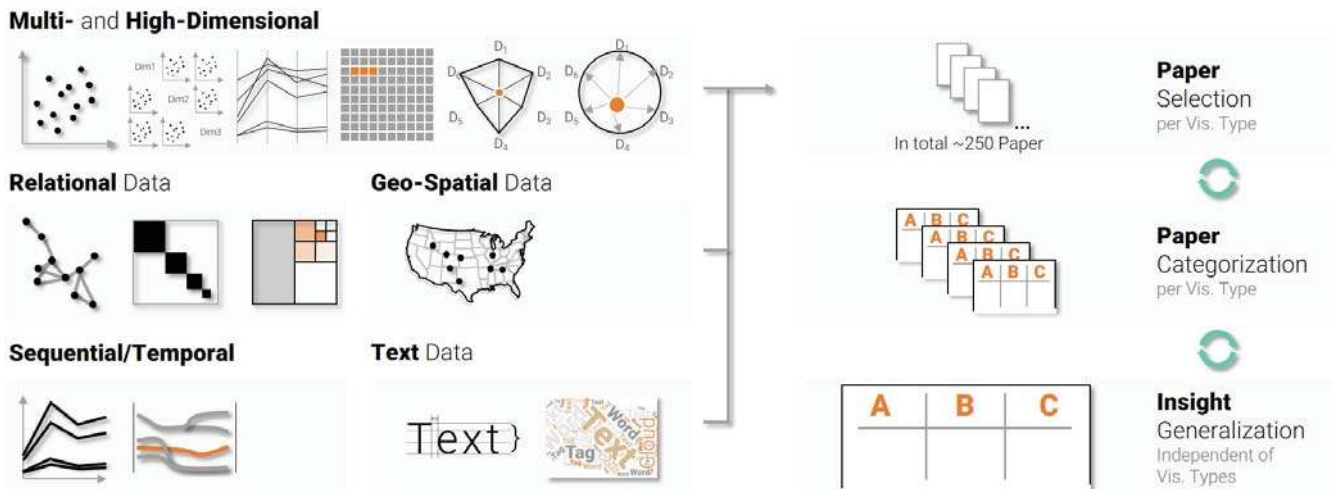
Figure 7: *Based on a comparative summarization approach, inspired by the grounded theory [SC94] and the structured content analysis theory [May00], we iteratively collected 134 presenting QM approaches from 14 distinct visualization fields acknowledging the need for QM. We condensed and aggregated our work in each QM subfield to derive higher-level findings reported in Background and Conceptualization.*

5. Methodology and Structure

Our surveying methodology is based on an iterative and comparative summarization approach, inspired by the grounded theory analysis of Strauss and Corbin [SC94] and the structured content analysis theory of Mayring [May00].

While a broad variance of options to structure the quality metrics field exist –we outline for example the historical perspective of this survey online http://visualquality.dbvis.de/history– we decided on a data-type driven approach.

Consistently, numerous books in the information visualization community agree that the primary data types to be visualized are a pure or mixed form of multi- or high-dimensional-, relational-, (geo-) spatial, temporal, or textual data [KAF*08, WGK10, Mun14]. Our survey borrows from this structuring of the field and summarizes for each data type the most prominent visualization types, such as Node-Link Diagrams or Matrix Representations for relational data, Parallel Coordinates, or Pixel-based Techniques for multi-dimensional data, or Scatter Plots and Scatter Plot Matrices for high-dimensional data. This enumeration is not exhaustive and explicitly targeted towards visualization subdomains that acknowledged a specific need for applying QMs.

We gathered an initial set of papers from an informal user study with domain experts (doctoral researchers and postdoctoral researchers with between 2–7 years of experience in respective visualization subdomains). Our paper selection was used to condense a set of high-level questions and evaluation criteria that guided in the following the expansion of the reference list by searching through the relevant visualization venues. Consequently, our survey should be seen as an educated selection of the concepts of quality metrics and does not claim comprehensiveness.

For each data type and visualization technique we base our analysis and the organization of each content section on a structured questionnaire, which incorporates the following aspects,

(i) **Visualization Description** outlines the basic concept of a specific visualization type, its primary purpose, its inherent constraints and requirements.

(ii) **Why do we need QMs?** motivates the use of QMs in this context, describing the perceptual/analytical benefits, sketches (computational) challenges, and refers back to the visualization definition part influenced most by the QMs.

(iii) **Typical Analysis Tasks** outlines analysis scenarios for the respective visualization and mentions how QMs can improve efficiency and effectiveness.

(iv) **Summary of Approaches** presents an overview of the influential QM work in the literature.

(v) **Evaluations Methods** shines a light on the evaluation approaches for QM-enhanced visualizations.

(vi) **Open Research Questions** summarize the future challenges with respect to the visualization design and states how QMs could be applied to overcome these problems.

In order to come up with a structured and valid abstraction of the field, we reported all of our findings in a table format (encoding phase), which can also be found online at http://visualquality.dbvis.de/summary. We iterated on the table results for consistency and developed iteratively a more and more refined view of the landscape. The core findings of these iterations are reflected and abstracted in the background sections, while specifics are highlighted in the respective subsections.

In total, we collected for this survey 134 papers from the various information visualization subfields. While our coverage is not exhaustive and biased toward impactful publications illustrating the fundamental concepts of this field, our goal is to provide a central document where concepts for multiple visualization types are defined and related, algorithms grouped into broader categories and discussed in contrast to each other, and, finally, we give an overview, of how quality metrics are systematically evaluated.

Multi- and High-Dimensional

Quality Metrics	Scatter Plots, SPLOMs	Albuquerque et al. [AEL*09], Aupetit et al. [AS16], Bertini et al. [BS04, BS05, BTK11], Ellis and Dix [ED07], Fink et al. [FHSW13], Friedman and Tukey [FT74], Matute et al. [MTL18], Micallef et al. [MPOW17], Lehmann et al. [LAE*12], Pandey et al. [PKF*16], Sarikaya et al. [SG18], Sedlmair et al. [STMT12], Shao et al. [SSB*16], Sher et al. [SBLC17], Sips et al. [SNLH09], Tatu et al. [TAE*11, TAE*09, TBB*10], Wilkinson et al. [WAG05],
	Parallel Coordinates	Ankerst et al. [ABK98], Claessen and Van Wijk [CVW11], Cui et al. [CWRY06], Dasgupta and Kosara [DK10], Ellis and Dix [ED07, ED06a, ED06b], Ferdosi et al. [FR11], Forsell and Johansson [FJ07], Heinrich et al. [HLKW12, HW13], Holten and Van Wijk [HVW10], Hurley and Oldford [HO12], Johansson et al. [JC08, JFLC08, JJ09, JFC14, JF16], Lind et al. [LJC09], Lu et al. [LHZ16], Netzel et al. [NVE*17], Palmas et al. [PBO*14], Peltonen and Lin [PL17], Peng et al. [PWR04], Rosenbaum et al. [RZH12], Tatu et al. [TAE*09], Walker et al. [WLP*13], Yang et al. [YPWR03]
	Pixel-based Vis.	Keim [Kei00], Schneidewind et al. [SSK06], Albuquerque et al. [AEL*10]
	Radial Vis.	Albuquerque et al. [AEL*10], Di Caro et al. [DCFMFM10]
	Glyphs	Borg and Staufenbiel [BS92], Fuchs et al. [FFM*13, FIB*14], Klippel et al. [KHLW09, KHW09], Peng et al. [PWR04], Yang et al. [YPWR03]

Relational Data

	Node-Link Diagrams	Adamec and Nešetřil [AN02], Battista G et al. [BGETT99], Dunne et al. [DRSM15], Huang et al. [HEH09], Nguyen et al. [NHEM17], Purchase [Pur97], Ware et al. [WPCM02], Wang et al. [WWS*18]
	Matrices	Behrisch et al. [BBH*17, BBR*16], Caraux and Pinloche [CP05], Chen [Che02], Ghoniem et al. [GFC04], Hahsler et al. [HHB08], Koren and Harel [KH02], Lee et al. [LPP*06], McCormick et al. [MDMS69, MSW72], Mueller et al. [Mue04, MML07b, MML07a], Niermann [Nie05], Petit [Pet03], Robinson [Rob51], Wilkinson [Wil05], Wilkinson and Friendly [WF09]
	Treemaps	Balzer and Deussen [BD05], Ghoniem et al. [GCB*15], Bederson et al. [BSW02], Bethge et al. [BHD17], Görtler et al. [GSWD18], Hahn et al. [HTMD14], Kong et al. [KHA10], Shneiderman [Shn92], Sondag et al. [SSV18], Tak and Cockburn [TC13], Tu and Shen [TS07], Ghoniem et al. [GCB*15]

Geo-Spatial Data

	Maps	Alam et al. [AKV15], Bertini et al. [BGS07], Brewer et al. [BMPH97], Correll and Heer [CH17], Keim et al. [KPSN04, KPSN03], Monmonier [Mon14], Nusrat and Kobourov [NK16], Rosenholtz et al. [RLMJ05], Sips et al. [SKNP04], Snyder [Sny82], Tobler [Tob04], Ward et al. [WGK10], Wongsuphasawat [Won16], Agrawala [Agr02], Andrienko et al. [AAB*13]

Sequential and Temporal Data

	Line Charts	Aigner et al. [AMM*07], Byron and Wattenberg [BW08], Cleveland [Cle94, Cle93a, Cle93b], Guha and Cleveland [GC11], Han et al. [HWZ*16], Havre et al. [HHWN02], Heer and Agrawala [HA06], MacEachren [Mac95], Palmer [Pal99], Talbot et al. [TGH11], Wang et al. [WWZ*17]
	Stacked Graphs	Byron and Wattenberg [BW08], Cui et al. [CLT*11], Cui et al. [CLWW14], Liu et al. [LZP*09]

Text Data

	Typographic Vis.	Alexander et al. [ACS*17], Beier [Bei09], Beier and Larson [BL13], Brath and Banissi [BB14a, BB14b, BB15, BB16], Strobelt et al. [SOK*16]
	Tag Clouds	Bateman et al. [BGN08], Buchin et al. [BCL*16], Cui et al. [CWL*10], Halvey and Keane [HK07], Hearst and Rosner [HR08], Jänicke and Chen [JC10], Lohmann et al. [LZT09], Paulovich et al. [PTT*12], Rivadeneira et al. [RGMM07], Schrammel et al. [SLT09], Seifert et al. [SKK*08], Sinclair and Cardew-Hall [SCH08], Smith et al. [SLPS*17], Strobelt et al. [SSS*12], Wang et al. [WCB*18], Witkin [Wit97], Wu et al. [WPW*11]

Figure 8: *The taxonomy of the reviewed approaches. For each quality metric approach, the taxonomy reports first author, the year, and the corresponding bibliographic reference.*

6. Multi- and High-dimensional Data

Multi- and high-dimensional data is typically provided in a table-like format in which rows correspond to data records/objects, and columns to their dimensions, attributes, features, or descriptors. For example, consider a collection of cars (data objects) that are described by, e.g., their color, brand, and horsepower (dimensions). Often, these datasets comprise *combinations* of numerical, categorical, and complex types such as geo-locations, images, and texts. In the following, we restrict ourselves to quality measures for (combi-

nations of) numerical and categorical dimensions. Quality metrics for more complex types are described in Sections 7, 8, 9, and 10.

Visualizations for multi- and high-dimensional data face two major challenges that also influence the computation of quality metrics: (1) datasets with a mix of numerical and categorical dimensions make it difficult to compute relations between objects (e.g., similarity) which is one of the fundamental concepts in many metrics. (2) The outstanding characteristic of datasets with a large number of dimensions is the *curse of dimensionality* [Bel61]. A huge number

of dimensions increase the possible visual mappings and the arrangement of dimensions. Non-relevant, redundant, and conflicting dimensions may hide interesting patterns in a sea of noise. And, the number of dimensions highly influence the interpretability of similarity measures [BGRS99, HAK00].

In the remainder of this Section, we will use synonymously the term high-dimensional for multi-dimensional, and multivariate data. We will describe and categorize quality metrics for Scatter Plots (6.1) and Scatter Plot Matrices (6.2), Parallel Coordinates (6.3), Pixel-based Techniques (6.4), Radial Visualizations (6.5), and G-lyphs (6.6). For each of the techniques, we describe the challenges and necessity of quality metrics, what they intend to measure, and outline the analysis tasks for the respective visualization. Afterward, we summarize the approaches and show their typical evaluation procedure, and outline open research questions.

6.1. Scatter Plots

One possibility to visualize high-dimensional data is to use Scatter Plots and Scatter Plot Matrices respectively. A Scatter Plot presents the data distribution of typically two variables as x and y axis in a Cartesian coordinate view. The main advantage of this visual representation is that the readability of single data instances as well as data patterns are straightforward and easy to understand. Figure 4 illustrates three examples of data distributions by means of Scatter Plots. One can clearly see if two variables correlate, reveal clusters, or patterns. In order to represent datasets with multiple dimensions, a Scatter Plot Matrix arranges all pairwise combination of dimensions into a tabular form, ordered by dimensions (see: Section 6.2).
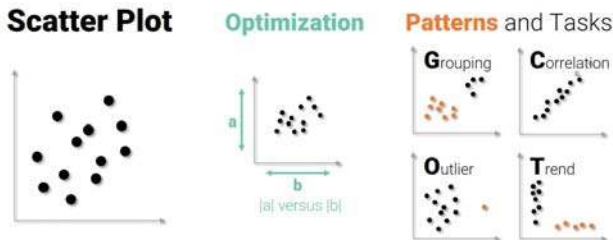

Figure 9: *Scatter Plots – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Scatter Plots?
Often, the input variables of a Scatter Plot can be displayed by different units, and the change of axis scalings may cause different patterns in the plot. To visualize clusters, patterns, and trends properly, the scaling of the two variables needs to be chosen carefully. Another well-known problem of Scatter Plot visualizations is to visualize large numbers of items, which often results in visual clutter. Visual clutter may obscure patterns in the data and makes it difficult for the user to find relationships among the dimensions. A challenge is to reduce the number of displayed elements but maintain the overall information at the same time. In recent years, several clutter reduction techniques have been developed to reduce the number of elements in a plot, which include sampling, filtering, clustering, and distortion techniques. However, each technique has its own requirements and objectives. Quality metrics and taxonomies may help find the best technique and settings for a given dataset or analysis task.

Typical Analysis Tasks for Scatter Plots
A Scatter Plot is used to investigate the relation between two different variables. It is useful to get a quick overview and helps indicate problems, unique properties, or anything interesting about the data. Interesting insights are, for instance, correlating variables, outliers, or meaningful patterns (e.g., regression models, trends, well-separated clusters). Sarikaya and Gleicher [SG18] presented a taxonomy of twelve low-level analysis tasks that support the analysis in Scatter Plot views. The defined analysis tasks are: identify objects, locate objects, verify objects, search for known motifs, browse data, identify outliers, characterize distribution, identify a correlation, explore neighborhood, numerosity comparison, object comparison, and understand distances.

In more advanced analysis scenarios, dimension reduction techniques are often used to map high-dimensional features into 2D projection views [WFC*18]. For instance, principal component analysis is a projection technique that uses traditional Scatter Plots to map high-dimensional data into a lower-dimensional space [WEG87].

Summary of Approaches.
Our ability to perceive patterns and trends in Scatter Plots is highly influenced by the aspect ratio. Cleveland [Cle93a] invented the principle called *banking to 45°*, which uses a midangle of 45° to enhance slope judgment for bivariate graphs (c.f. Section 9.1). Applied to Scatter Plots, an improved aspect ratio selection, such as with the banking to 45° quality criterion, can be applied to emphasize trends in the dataset [Cle93a]. This relationship between task (trend detection) and quality criterion (aspect ratio) was also examined and validated by Fink et al. in [FHSW13]. In addition, Fink et al. [FHSW13] uses Delaunay triangulation to generate Scatter Plot projections and measure the quality by calculating a small total edge length or large minimum angle of the triangles.

By assuming that the aspect ratio is chosen well, there still remains the question if the visual representation is appropriate for the data or not. A taxonomy of different visual factors to separate clusters in Scatter Plot well was given by Sedlmair et al. [STMT12]. The presented taxonomy is based on classified data and considers *within-class* and *between-class* factors to guide design and evaluations of cluster separation measures. Furthermore, clutter must be considered to present point distributions clearly. An overview of different clutter reduction techniques including benefits and losses for Scatter Plot visualizations is given by Ellis and Dix [ED07]. Regarding quality metrics for clutter reduction, Bertini and Santucci [BS04, BS05] proposed a feature preservation approach to improving visual perception of 2D Scatter Plots. Their metric includes an automatic sampling strategy based on a perceptual user study to find an appropriate sampling ratio.

Evaluation Methods for Scatter Plot Quality Metrics.
To assess improvements in visual perception, user studies were often conducted [LHT15, PKF*16]. Micallef et al. [MPOW17] implemented several models and metrics of human perception in a cost function to improve the visual low-level perception of a Scatter Plot. Based on input data and task, an optimizer automatically enhances design parameters such as marker size and opacity, aspect ratio. Various Scatter Plot design choices were investigated in [SG18] based on data characteristics and analysis tasks.

Open Research Questions.
The traditional Scatter Plot is a well-known visualization technique

and has been further developed over the last decades. Today, various kinds of visual optimizations and data preprocessing techniques exist to improve the final representation. For instance, there are density based modifications or combined representations to increase the information content on the visualization. However, what is missing in the literature are state-of-the-art reports about Scatter Plot related techniques and optimizations.

6.2. Scatter Plot Matrices

A Scatter Plot only visualizes the relationship between two dimensions (bivariate data). To investigate the whole data space of a high-dimensional data set a Scatter Plot Matrix (SPLOM) can be used, which shows all pairwise Scatter Plots of the different variables n in a matrix. A SPLOM consists of n^2 cells, where each column and row reflects one data dimension. Thus, data analysts can inspect the changes of independent variables according to a dependent variable by scanning the rows, respectively columns, of the matrix. Figure 10 shows a basic SPLOM and typical interesting patterns for analysis. Hence, we obtain a SPLOM including $n^2 - n$ single Scatter Plots.
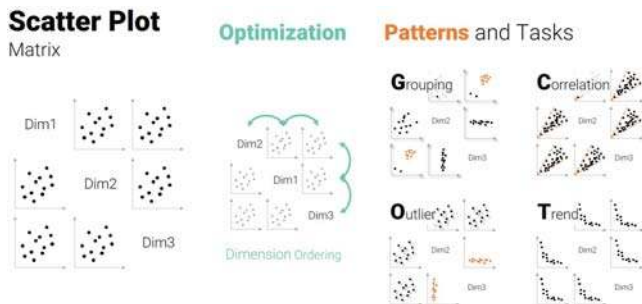

Figure 10: *Scatter Plot Matrix – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Scatter Plot Matrices?
Exploratory data analysis in large Scatter Plot Matrices is a challenging task, since the number of projection views grows quadratically with the number of dimensions. Furthermore, the goal of exploration is based on a given analysis task or user, and typically not all Scatter Plot views are potentially relevant. Thus, often a manual exploration for finding interesting patterns, trends or clusters becomes exhausting and ineffective. To improve the exploration, quality metrics can be used that apply computational measures based on data and/or image space to identify the most interesting views according to the analysis tasks.

Typical Analysis Tasks for Scatter Plot Matrices
A SPLOM is often used to get an overview of all bivariate correlations (via Scatter Plots) in a higher dimensional data space. This is particularly helpful to identify specific variables that might have similar patterns across various dimensions, e.g., correlation, classification, clusters, or trends.

Due to the orthogonal pairwise projections of dimensions in a SPLOM, a horizontal or vertical exploration enables the investigation of data transformations by exchanging one dimension. For example, a column-wise exploration allows the user to discover transformations by exchanging the independent variable and a row-wise exploration by exchanging the dependent variable. Shao et al. [SSB*16] used color coding in combination with a motif-based

dictionary to highlight column-wise and row-wise coherence of segmented patterns in s SPLOM. This work also encourages to take the investigation of local patterns into the analysis process and focus on interest measures derived from local motifs in the data.

Furthermore, SPLOM-like representations are suitable for subspace analysis tasks, such as finding clusters or interesting subspaces. Yuan et al. [YRWG13] used a dimension projection matrix in which rows and columns represent multiple dimensions and the Scatter Plots are based on dimension projection.

Basically, all low-level perception task for single Scatter Plots (mentioned in Section 6.1) can be applied to a larger projection space. For the analysis in SPLOMs, these tasks are usually extended to a comparison task among multiple Scatter Plots (mid-level perception task). Sarikaya and Gleicher [SG18] derived twelve basic analysis tasks that are supported in Scatter Plot and SPLOMs respectively (c.f., Typical Analysis Tasks for Scatter Plots in Section 6.1).

Summary of Approaches.
In data analysis, methods for mapping multivariate data into lower dimensional space have been used for many decades [KW78, WEG87]. However, one of the major problems of these mappings is that the resulting outcome is often difficult to interpret. One influential approach by Friedman and Tukey [FT74] that tackles this issue is called Projection Pursuit. Projection Pursuit is a linear mapping algorithm that uses interpoint distances and the variance of point swarm to pursue optimum projections. Later, Tukey and Tukey [TT85] invented an exploratory visualization method for SPLOMs (Scagnostics). Wilkinson et al. [WAG05] followed up on their research and introduced graph-theoretic measures for computing scagnostic for large datasets. The method is based on proximity graphs and extracts nine characteristics that describe the point distributions of the Scatter Plot space. It has been shown that Scagnostics can serve for many applications and help to detect anomalies in time series, find specific patterns or sort large SPLOMs [WW08, DAW13, DW14a, DW14b].

Another common approach to index the interestingness of Scatter Plots is to consider the class consistency information of labeled points. For instance, Sips et al. [SNLH09] propose two quantitative measure of class consistency, one based on the distance to the class's center of gravity (distance consistency), and another based on the entropies of the spatial distributions of classes (*distribution consistency*). Tatu et al. [TAE*09] used similar ranking measures based on the image space of the Scatter Plot visualization to identify potentially relevant structures. For unclassified data, they used a *rotating variance measure* (RVM) to find linear and non-linear correlations in the data. For classified data, they measure the overlap between the different classes and rank Scatter Plots that show well-separated classes the best (*class density measure*). Later, Tatu et al. [TAE*11] extended their class density measure and introduced *class separating measure* to control the balance between the property of separation and dense clustering. A recent work of Matute et al. [MTL18] showed that a skeleton-based metric including shape and orientation information outperforms *RVM* and *Scagnostics* in perceptually-based similarity.

Albuquerque et al. [AEL*09] utilized the aforementioned quality measures for a quality-aware sorting of SPLOMs, the so-called class-based Scatter Plot matrix (C-SPLOM). Lehmann et al. [LAE*12] introduced another visualization scheme including detail-on-demand

interactions that produces an abstract and interpretable SPLOM (A-SPLOM) by using known quality measures.

A current approach by Shao et al. [SSB*16] measures the interestingness by taking frequency properties of similar local patterns into account. The approach applies a Bag-of-Visual-Words concept that considers local motifs as visual word and ranks the interestingness based on the number of interesting motifs in a plot. Moreover, an extensive survey of quality metrics for Scatter Plot and other visualization techniques were carried out by Bertini et al. [BTK11]

Evaluation Methods for Scatter Plot Matrix Quality Metrics.
The evaluation of quality metrics has gained increasing importance in the recent years. These works either focus on evaluating the connection between human perception and quality metrics (effectiveness) or demonstrating the usefulness of quality metrics base on various use case scenarios (efficiency).

For example, Projection Pursuit is demonstrated by various experiments on artificial and research data. Wilkinson et al. [WAG05, WW08, DAW13, DW14a, DW14b] evaluated the performance and usefulness of their Scagnostics tools by showing use cases and experimental results on different datasets. Actually, most approaches are evaluated by use cases and demonstrate the benefits by various scenarios [TAE*09, AEL*09, TAE*11, SSB*16]. For instance, the class consistency measures by Sips et al. [SNLH09] were applied to synthetic data and various well-known data sets from the UCI repository [DKT17]. Classified data are ranked according to how consistently the high-dimensional classes are embedded in the 2D projection space. For unclassified data, a clustering algorithm is applied to generate high-dimensional class structures.

More recent work by Lehmann et al. [LHT15] and Pandey et al. [PKF*16] have investigated the human perception on Scatter Plot patterns and present both crowdsourcing studies to prove evidence that this connection exists. Sedlmair and Aupetit [SA15] even present a data-driven framework for quality measure evaluation. Their approach tries to mitigate the impact of (relative) human judgments by relying entirely on ground-truth data. However, this in turn also indirectly implies some sort of user involvement. By using this framework, Aupetit and Sedlmair [AS16] evaluated a large number of visual separation measures for pre-classified data. They systematically generated 2002 visual separation measures by combining neighborhood graphs and class purity function with different parameterizations. As a result, they identified measures that outperforms the distance consistency measure. Sher et al. [SBLC17] conducted a study about the human perception of correlations in Scatter Plots. Their study reveals that humans perceive correlations differently compared to the statistical measure of Pearson's product-moment correlation coefficient.

Bertini et al. [BTK11] pointed out that all quality metrics that work in the image space try to simulate the human pattern recognition machinery and therefore, it is needed to validate and tune the metrics in a way that the parameters take models of human perception into account. Together with other colleagues [TBB*10], they presented a user study about human perception and quality metrics, where they compared the outcome of quality metrics with human rankings. The usefulness of Lehmann's A-SPLOM [LAE*12] was evaluated by a controlled experiment including 12 participants. The task of the study was to select relevant plots from different SPLOM configurations (A-SPLOM, unsorted A-SPLOM, sorted SPLOM,

unsorted SPLOM). Finally, they compared mean and variance values of the number of selected plots to the values of the quality measures.

Open Research Questions.
Even though a lot has been done in the field of quality metrics for Scatter Plot visualization, there are still some directions that can be further investigated. One possible direction could be the integration of human sensing technologies, e.g., eye tracking or motion tracking, to investigate the behavior of users during an analysis task. For instance, prior research of Shao et al. [SSES17] has shown that eye tracking devices can be used to track already explored patterns, and thus support the exploration of varying patterns in large Scatter Plot spaces. Furthermore, eye tracking has also been used the evaluation of Scatter Plots and Parallel Coordinates. By using an area-of-interest (AOI) approach, Netzel et al. [NVE*17] showed how participants act during analysis tasks and identified different reading strategies. Consequently, these sensing measurements could be integrated into the quality metrics-driven visual analytics pipeline and enrich the quality criterion inputs (user U, task T).

6.3. Parallel Coordinates

Parallel Coordinates [Ins09] are one of the most popular visualizations for multi- and high-dimensional data. Introduced to the information visualization community by Inselberg [Ins85], the technique gained popularity by enabling analysts to explore patterns across a large set of dimensions. Equally-spaced vertical axis represent the dimensions of the dataset; the top of the axis corresponds to the highest, the bottom to the lowest value in each dimension. Data points are mapped to polylines across the axis, such that the intersection between an axis and a polyline marks the data value. This visual mapping allows analysts to spot high-level patterns, as well as single data points of interest.
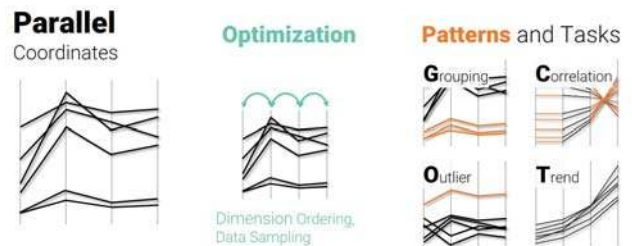

Figure 11: *Parallel Coordinate Plots – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Parallel Coordinates?
Parallel Coordinates face three major challenges: (1) With an increasing number of data records, the patterns start to disappear due to *overplotted* lines [ED07]. (2) A perceived pattern in Parallel Coordinates depends primarily on the ordering of the dimension axis [JJ09]. A proper ordering (for a specific task) can reveal unknown patterns while a non-useful ordering may hide them. Consider the example in Figure 12. Two different ordering strategies are applied to reveal clusters and correlations. (3) A large number of dimensions decreases the available screen space between two axes and results in cluttered plots; in particular when many data records are shown [DK10]. (4) The perception of positive and negative correlation is not symmetric: negative correlations are visible more clearly [HLKW12, LMVW10].
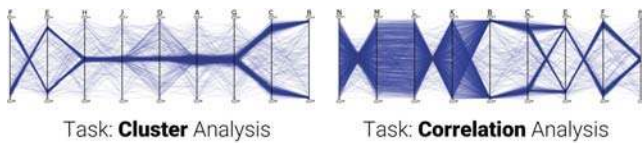
Figure 12: *Two partially different subsets of dimensions with axis ordering strategies optimized for cluster analysis (left) and correlation analysis (right). Figure adapted from [JJ09].*

Quality metrics need to tackle these challenges by measuring the quality of a perceived pattern and the amount of clutter/overplotting in order to be able to guide ordering and sampling strategies. However, there are $n!$ possible dimension permutations (based on the assumption that we plot every dimension exactly once). Having a quality criterion that measures the quality of one particular permutation, Ankerst et al. [ABK98] prove that finding the optimal ordering can be reduced to the traveling salesman problem and is therefore $\mathcal{NP}$-complete. As a consequence, not only quality criteria but also efficient optimization algorithms are necessary.

Typical Analysis Tasks for Parallel Coordinates
Countless applications from various (research) domains have been tackled with Parallel Coordinates. In a recent state-of-the-art report by Heinrich and Weiskopf [HW13], the tasks of these applications are categorized according to the established KDD taxonomy by Fayyad et al. [FPSS96]: *classification*, *regression*, *clustering*, *summarization*, *dependency-modeling*, and *change and deviation detection*. In analogy, we show in Figure 11 four of the main visual patterns for that help to accomplish these tasks: *grouping*, *correlation*, *outlier*, and *trend*. Quality metrics should be able to re-order and de-clutter Parallel Coordinates such that these patterns are visible to the analyst (based on the current analysis task).

Summary of Approaches.
A multitude of quality metrics has been presented for Parallel Coordinate Plots. The approaches can be separated into *quality criteria* measuring the quality of one *visualization definition* and *optimization algorithms* that optimize the adjustable parameters. In the following, we will first describe and discuss the variety of quality criteria, followed by the applied optimization algorithms.

The first criterion described in the literature has been developed by Ankerst et al. [ABK98] with the argumentation that a similarity-based ordering will reduce visual clutter. In their data space approach, the authors propose finding a perceptually "good" ordering by measuring the Euclidean distance between two dimensions on a global level, or by partial similarity based on a defined threshold. The quality criterion measures the sum of distances between all neighboring dimensions, which needs to be minimized by the optimization algorithm. Yang et al. [YPWR03] extend the idea by applying a clustering on the dimensions first. Due to the resulting hierarchy, the search space of permutations can be reduced by considering only dimensions within one cluster.

Another similarity-based method is proposed by Peng et al. [PWR04]. The authors claim that the source of clutter can be caused by distortions of the data distribution, e.g. due to outliers. Peng et al. define an outlier based on the nearest neighbor algorithm and propose a quality criterion based on the proportion of outliers between two neighboring dimensions. Similar to the previous approaches,

clutter is only measured between two neighboring dimensions in the visualization.

A quality criterion for supporting nearest neighbor searches is proposed by Peltonen and Lin [PL17]. In their approach, the similarity between axes is computed using the Kullback-Leibler divergence of probabilistic neighborhood distributions.

Ellis and Dix [ED06a, ED06b] propose three methods to estimate the occlusion of lines in Parallel Coordinates: (1) overplotted% (percentage of pixels with more than one *plotted* point), (2) overcrowded% (percentage of pixels with more than one *existing* point), and (3) hidden% (percentage of plotted points *hidden* due to overplotting). Ellis and Dix propose several data space algorithms to count the number of pixels or points respectively. All criteria can be applied globally or in areas of interest, e.g., by a sampling lens [EBD05].

Several methods quantify the difference in the data distribution between the original space and a subset of data records or dimensions. Cui et al. [CWRY06] measure the difference of data density for all dimensions using a histogram approach. The quality criterion retrieves the difference between the histogram of the data sample and the histogram of the original data. In the same paper, the authors extend the idea by quantifying the similarity of each record in the original space with its nearest neighbor in the sample. An image space method by Johansson and Cooper [JC08] transforms the visualization into a so-called distance map [RP66] in which each pixel describes the distance to its closest object. The quality criterion measures the similarity between the distance maps of the original and the sampled data.

Several approaches argue that the first dimension attracts the most attention of the user. Therefore, it should be considered in the ordering. Lu et al. [LHZ16] use Singular Value Decomposition to measure the contribution of each dimension to the data space. Highly contributing dimensions are sorted up front. Yang et al. [YPWR03] consider the importance of a dimension (e.g., by variance) in their similarity-based ordering. A different method is proposed by Ferdosi and Roerdink [FR11] to promote grouping patterns. They do not only consider a pair-wise combination of dimensions, but rather search for high(er)-dimensional structures by using a subspace clustering algorithm. The quality of a subspace is measured by the density distribution [FBT*10] and an implicit algorithm sorts the dimensions based on the quality of each individual subspaces.

Tatu et al. [TAE*09] introduce three image-based quality criteria to measure the quality of perceived clusters. The assumption of all methods is that clusters are usually represented by clustered lines with a similar position and direction. The image of the visualization is transformed by a Hough transformation [VC62] into a new image, such that lines with a similar slope and interception are at a close location. The quality criterion measures the clusteredness between two dimensions within the Hough space. For datasets with given cluster labels, Tatu et al. adapt their measure and focus on (1) the intra-class similarity, and (2) cluster overlap by measuring the difference between the Hough space images per cluster.

One of the most central image-based QM approach is proposed by Dasgupta and Kosara [DK10]. Pargnostics, following idea of Scagnostics [WAG05] for Scatter Plots, are a set of seven quality criteria for Parallel Coordinates: *number of line crossings*, *angles of crossing*, *parallelism*, *mutual information* (dependency between

variables), *convergence* and *divergence*, *overplotting*, and *pixel-based entropy* (randomness = uncertainty). The proposed measures are computed from 1D statistics and 1D/2D distance histograms, which allow for a rapid computation. The optimization algorithm can make use of a weighted combination of features.

Finally, Johansson and Johansson [JJ09] provide an interactive analysis of the whole high-dimensional dataset based on different quality metrics that can be selected and weighted by the user. The authors describe three criteria to measure the quality of a plot: (1) *Correlation analysis* by the Pearson correlation coefficient [LRN88] between neighboring dimensions. The quality scores between neighboring dimensions are aggregated for the entire plot. (2) *Outlier detection* based on a grid-based density computation. The quality criterion combines the number of dimensions and the distance to the nearest neighbor across multiple dimensions. (3) *Cluster detection* by a subspace clustering approach (e.g., Mafia algorithm [NGC01]). For each subspace cluster, a quality score is computed representing density, dimensionality, and the fraction of the covered dataset.

So far, we have discussed quality criteria for combinations of two or more dimensions. In order to find an optimal ordering for the entire Parallel Coordinates Plot, optimization algorithms are necessary. As shown by Ankerst et al. [ABK98], the reordering task in Parallel Coordinates is $\mathcal{NP}$-complete. The literature does not provide any novel algorithmic solutions, but rather applies existing approaches. To name a few: heuristic algorithms are used in [JJ09], a genetic approach is presented in [ABK98], and graph-based algorithm is applied in [TAE*09, DK10, HO12].

Evaluation Methods for Parallel Coordinates Quality Metrics.
New quality metrics are mostly evaluated by showing examples based on synthetic or real-world datasets. Often, the performance of optimization algorithms is depicted in terms of efficiency. Only a few approaches compare multiple quality criteria: Ellis and Dix [ED06b] and Cui et al. [CWRY06] *empirically* compared their own approaches with each other. Ferdosi and Roerdink [FR11] systematically compared their subspace clustering approach with the similarity clustering method of Ankerst et al. [ABK98], the clutter-based method of Peng et al. [PWR04], the Hough space method by Tatu et al. [TAE*09], and the hierarchical dimension clustering method by Yang et al. [YPWR03].

The number of user-centered evaluations, investigating the perceptual aspects of QMs, is limited, as also discussed by a recent survey of Johansson and Forsell [JF16]. Few studies exist to measure the influence of clutter: Holten and van Wijk [HVW10], Heinrich et al. [HLKW12] and Palmas et al. [PBO*14] quantitatively analyze the reduction of clutter through edge bundling techniques or different variations of Parallel Coordinates extensions. Rosenbaum et al. [RZH12] evaluates the readability of Parallel Coordinates under different densities of data points, but focuses on progressive analytics argumentations. A qualitative and quantitative evaluation scheme considering the ordering of dimensions is discussed in Claessen and van Wijk [CVW11] and Walker et al. [WLP*13]. Further studies exist to compare axis arrangements in 2D vs. 3D Parallel Coordinates: Forsell and Johansson [FJ07], Johansson et al. [JFC14, JFLC08], Lind et al. [LJC09]. A recent eye-tracking study of Netzel et al. [NVE*17] revealed that participants pay more attention towards the center of Parallel Coordinates Plots. This finding

stands in contrast to the dimension contribution-based approaches, for example, by Lu et al. [LHZ16].

Open Research Questions.
Promising future work is summarized by the survey of Johansson and Forsell [JF16]. The community has developed many quality metrics for Parallel Coordinates. The concepts of useful orderings and clutter reduction approaches for the underlying method differ significantly. However, there are no user studies that compare the different metrics for different tasks and different data characteristics. Based on such findings, the community could further develop task-dependent quality metrics that support the perception of humans.

6.4. Pixel-based Techniques

Pixel-based techniques create a separate view (called *subwindow*) for every dimension of a dataset. Within each subwindow, every data record is mapped to exactly one pixel, colored according to the value in the respective dimension [Kei00]. Pixel-oriented visualizations do not face overplotting issues, and they are designed to display large amounts of data without aggregation. The number of data points to be visualized is only limited by the available screen space.
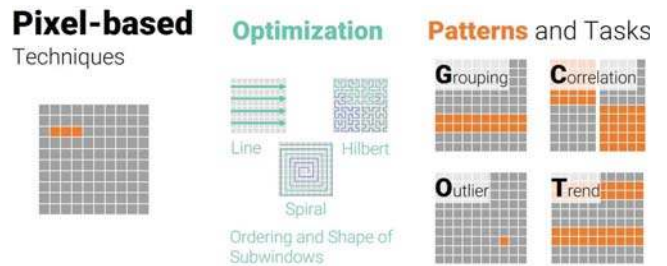

Figure 13: *Pixel-based Visualizations – Optimization Goals, Analysis Tasks & Visual Patterns*

The most important aspect is the layout of pixels within each subwindow. For each window, the same layout is applied in order to make the dimensions comparable. Generally, the data points require an ordering, such as a natural order (e.g., by time or size), or the result of a function (e.g., order of nearest neighbors to a query object). Design recommendations by Keim [Kei00] and Wattenberg [Wat05] propose that data points need to be layouted such that the given ordering of the data is approximated in the subwindows. This means, data points that are nearby in the ordering, should end up nearby in the visualization. For rectangular-shaped subwindows, space-filling curves are proposed to optimize these recommendations, for example Hilbert Curves [Hil91] or H-Curves [NRS97].

Why Do We Need Quality Metrics for Pixel-based Techniques?
Pixel-based visualizations are designed to display large amounts of data, but only for individual attributes. With high-dimensional data with thousands of attributes being more and more common, it is practically impossible to manually inspect the visualization for each individual attribute for interesting patterns. Hence the need for quality metrics for pixel-based visualization techniques. They help users to analyze high-dimensional data sets by calculating a quality metric for each attribute. These quality metrics can be used to identify interesting attributes. According to Keim [Kei00], there are four properties that have to be considered when designing pixel-oriented
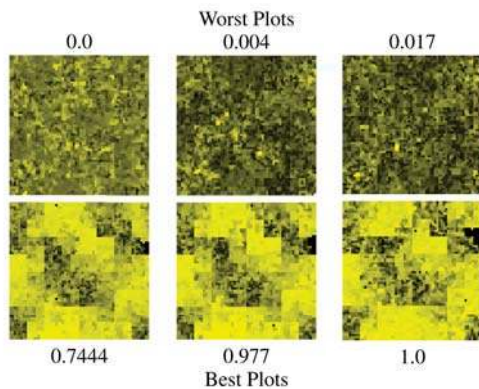
Figure 14: *Jigsaw maps and quality metric scores of the Ozone datasets created by Albuquerque et al. [AEL*10]. Depicted are six Jigsaw maps along with their calculated Noise Dissimilariy Measure. The top row shows the worst three plots and the bottom row the best three plots with clearly visible patterns.*

visualizations. The color mapping, the arrangement of pixels, the shape of the subwindows, and the ordering of the dimensions. For each of these properties, Keim [Kei00] presents *design recommendations*. For instance, the usage of space-filling curves like the Morton curve [Mor66] for the arrangement of pixels. The problem is that the methods proposed by Keim such as the ordering of dimensions, the shape of the subimages, and arrangement of pixels require solving complex optimization problems. Some of which are proven to be $\mathcal{NP}$-hard [ABK98].

Typical Analysis Tasks for Pixel-based Techniques

Pixel-based visualization techniques are useful for solving four different tasks on large high-dimensional data, as depicted in Figure 13. When analyzing a single dimension, pixel-based visualizations can be used to identify clusters and outliers. Clusters, such as visible in Figure 14, can be identified by finding local regions of similar color. Outliers, in contrast, are depicted as points with outstanding colors in comparison to their surrounding region. Trends are depicted by consistently reoccurring occurrences of similar color spread out over the pixel plot. When considering multiple dimensions, pixel-oriented visualizations can be used to identify correlations between different dimensions. If a cluster occurs in multiple dimensions, this can be an indication for a positive correlation, if they share a color, or negative correlation, if they consistently depict a different color. However, finding these visual patterns is only possible if the ordering between and within dimensions is done appropriately.

Summary of Approaches.

The existing approaches for pixel-based visualizations can be divided into data space, image space and hybrid approaches. Keim [Kei00], in addition to his general optimization algorithms for pixel-based visualizations, presents such data space quality criteria for geospatially-related data. The presented quality criteria focus on the layout and positioning of the pixels in the resulting visualization and measure, for instance, the position-preservation of the layout algorithms, the relative position-preservation or the relative distance-preservation.

In addition to these data space approaches, also two image space approaches, Pixnostics [SSK06] and the Noise Dissimilarity

Measure (NDM) [AEL*10] were presented for pixel-based displays. Pixnostics calculates the information content of a pixel-based visualization by calculating either the entropy or the standard deviation on the distribution of gray-level histograms in different grid cells. If the calculated score for a gray-level histogram of a cell is between two user-defined thresholds, it is considered to be interesting. However, this requires a manual setting of the interestingness thresholds. NDM uses the dissimilarity between a visualization and a noise image generated by a random permutation of the original visualization. Since the characteristic of the noise image is supposed to be the total absence of structure, visualizations with a large Noise Dissimilarity Measure are considered to have a higher potential relevance, as shown in Figure 14.

Evaluation Methods for Pixel-based Quality Metrics.

In the presented works, there is no standard evaluation technique for pixel-based visualization techniques apparent. Keim [Kei00] provides a quantitative evaluation for geospatially-related data by measuring and comparing the position and distance preservation of different layout algorithms. Schneidewind et al. [SSK06] and Albuquerque et al. [AEL*10] both show the effectiveness of their quality metrics in a use case study, by showing the potential to find interesting visualization. Both start with a set of pixel-based visualization. Schneidewind et al. create their test set by randomly permuting the pixels of a Jigsaw map and Albuquerque et al. create multiple visualizations for the Ozone dataset, as shown in Figure 14. Both use their respective quality metric to calculate a score for each visualization in their test set and show that visualizations with a high QM score contain interesting visual patterns.

Open Research Questions.

The quality metrics for geospatially-related data proposed by Keim [Kei00] focus on the , e.g. by calculating the position-preservation of the resulting visualization, while the quality metrics proposed by Schneidewind et al. [SSK06] and Albuquerque et al. [AEL*10] focus on the image space. Both of these image space approaches are particularly useful for identifying groupings. However, for other analysis tasks the pixel-visualization suitable quality metrics are missing. A beneficial line of research could seek to adapt e.g., TreeMaps QM (see: Section 7.3) to this domain. Furthermore, a comparative and user-agnostic evaluation of the existing approaches could help to identify a baseline for further research directions.

6.5. Radial Visualizations

Radial visualizations for high-dimensional data arrange the data in a circular or elliptical fashion. Draper et al. [DLR09] presents a general survey on the topic referencing 88 works. Prominent techniques for high-dimensional data include, but are not limited to: the *MoireGraph* [JM03], the *TimeWheel* and *MultiComb* visualization both proposed by Tominski et al. in [TAS04] and the projection-based techniques, such as RadViz [HGM*97] and Star Coordinates [Kan00]. Note that visualizations, such as Pie Charts, Sunburst, or Radar Charts, albeit being radial visualization are explicitly excluded here, since their optimization focuses on storytelling and semantic aspects, c.f., high-level quality metrics in Section 3.2. The development of perceptual quality metrics was mainly driven by high-dimensional (projection-based) radial visualizations and thus will be the focus of this Section.
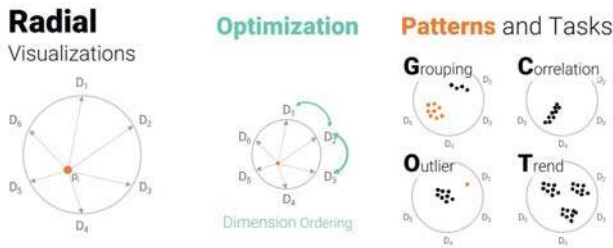
Figure 15: *Projection-based Radial Visualizations – Optimization Goals, Analysis Tasks & Visual Patterns*

Projection-based radial visualizations are two-dimensional projections of high-dimensional data into a circle. For RadViz, the dimensions of a dataset X are represented as points that are evenly spread around the circumference of a circle. Each instance of the dataset $x_i \in X$ is also represented as a point inside the circle. The positioning of each instance x_i can be determined by connecting it with springs to each of the dimension representatives on the circumference of the circle. The final position of a point x_i is determined by the point p_i where the sum of all spring forces is zero and can be computed as $p_i = \frac{\sum_{j=1}^{n} d_j x_{i,j}}{\sum_{j=1}^{n} x_{i,j}}$, with d_j denoting the vector pointing from the center to the position of the respective dimension on the circumference [AEL*10]. Star coordinates apply a nearly identical mapping of points into a circle, but without the nonlinear normalization, for which the denominator of the previous equation is responsible. As a result of this, RadViz is especially advantageous for sparse data, but its nonlinearity may hamper several other exploratory analysis tasks [RSRDS16]. Rubio et al. point out, that due to their similarity, many algorithms, such as the quality metrics developed by Albuquerque et al. for RadViz, can be directly applied to Star Coordinates and vice versa.

Why Do We Need Quality Metrics for Radial Visualizations?
Similar to Parallel Coordinates Plots (see: Section 6.3), radial visualizations are highly dependent on the ordering of dimensions, which is in turn dependent on the user's task. For example, if one data instance has high values in two neighboring dimensions, it is plotted more closely to the circumference, in addition, another data instance with high values in two opposite dimensions is plotted more closely to the center of the circle. Given an ordering of the dimensions, quality metrics can help to identify if the resulting visualization has interesting patterns for a specific user task. However, finding a suitable ordering is one of the key problems, which Ankerst et al. [ABK98] proved to be $\mathcal{NP}$-complete. Hence, QMs for radial visualizations are not only necessary but also efficient techniques to explore the search space of possible dimension orderings.

Typical Analysis Tasks for Radial Visualizations
RadViz was first proposed by Hoffman et al. [HGM*97] to help with the classification of DNA sequences. In their work they compare visualizations of multi-dimensional DNA sequence data. They compare, on the one hand, visualization techniques which are able to display all dimensions, i.e., RadViz and Parallel Coordinates, to techniques which use dimension reduction techniques to produce 2-dimensional visualization, such as Sammon Plots, on the other hand. They conclude that, although that some patterns can still be seen in the dimension-reduction techniques, the exact symmetry is lost, which is an inherent problem of such techniques due to the difficulty

of choosing the important dimensions. RadViz can also be used to tackle various different tasks, as shown in Figure 15. Nováková and Štěpánková show how radial visualizations can be used to detect trends in time-series data [NŠ11]. Mrarmor et al. [MLDZ07] use radial visualizations for outlier detection in lung cancer data based on gene expressions of six genes. Finally, Bertini et al. [BDS05] show how an extension of RadViz can help to detect correlations in data. Kandogan advertises star coordinates as a means for cluster, trend and outlier detection likewise [Kan00].

Summary of Approaches.
Albuquerque et al. [AEL*10] show that due to the scatter properties of RadViz, most quality measures for Scatter Plots (see: Section 6.1), may be applied to RadViz as well, such as the *Class Density Measure* [TAE*09] for labeled datasets. They also introduce the *Cluster Density Measure (C_lDM)* as a new quality metric to rank visualizations based on how well-defined the clusters of the resulting projection are. This image-space based technique first applies an image clustering algorithm and then calculates the quality metric score based on the found cluster properties. They follow the following computational steps; calculate a density image based on the local neighborhood in the original visualization; smooth the density image by applying a Gaussian filter; identify clusters with the help of Laplace filters; and calculate the C_lDM measure, defined as:

$$C_lDM = \frac{1}{K} \sum_{k=1}^{K} \sum_{l=k+1}^{K} \frac{d_{k,l}^2}{r_k r_l}$$

where K is the number of detected clusters, $d_{k,l}$ the Euclidean distance between the cluster centers c_k and c_l and with r as the average radius of a cluster. Thus projection clusters with a small intra-cluster and large inter-cluster distance are assigned high values.

Another approach to calculating quality metrics for radial visualizations is presented by Di Caro et al. [DCFMFM10]. They determine the visual usefulness of a projection by using the Davies-Bouldin (*DB*) index [DB79]. The *DB* index is known to be one of the best methods to measure the inter- and intra-cluster separation. A smaller *DB* index represents more compact and separated clusters. However, if a high-dimensional dataset d has a high DB index, it may become difficult for the projected data p to offer a high-quality visualization. Thus, the DB index is not directly used as a quality metric, but rather the ratio R between the index of the high-dimensional data DB_d and the projected data DB_p is taken, with a high R corresponding to a higher visualization quality.

Evaluation Methods for Radial Visualization Quality Metrics.
Both quality metrics presented in the last section are used to evaluate new dimension-ordering techniques for RadViz. Di Caro et al. [DCFMFM10] provide an independent and a RadViz-dependent formalization of the dimension arrangement problem, which was formalized by Ankerst et al. [ABK98] in a generic context. They provide an exhaustive evaluation of both of these dimension arrangement techniques, partly evaluating the visual quality of the resulting arrangements. Moreover, Albuquerque et al. [AEL*10] propose a greedy RadViz generation algorithm in which they start with a two-dimensional RadViz and iteratively add the remaining dimensions by checking which dimension they have to add for optimizing a quality metric. Additionally, they provide three comparisons of the resulting visualizations, using the original RadViz algorithm, the

t-statistics algorithm of Sharko et al. [SGM08], and their algorithm, concluding that using their algorithm, the resulting projections show a better cluster separation.

Open Research Questions.
So far two algorithms were proposed to measure the visual quality of visualizations generated by RadViz. One data space and one image space technique. Both approaches have shown, that their quality metric can be used to determine the visual quality of a resulting projection and they can even be applied during the construction of RadViz visualizations. However, the shortcomings are that both of these techniques focus on only one aspect, the intra- and inter-cluster separation. As previously shown, there are various possible applications of RadViz, with grouping only being one of these applications. In future work, quality metric for these different tasks, such as outlier or trend detection, or, if possible, a general quality metric usable for various tasks should be developed. Additionally, as Rubio et al. pointed in their comparative study of RadViz and Star Coordinates [RSRDS16], algorithms designed for one technique, may be applied to the other. Therefore, when developing techniques for one technique, they recommend considering whether they would be appropriate for the other technique as well.

6.6. Glyphs
In the 1960s, the term glyph was used as a synonym for the metroglyph [And57]. However, over the years different glyph designs emerged and the initial definition was adapted to also describe new representations. As a result, the term glyph is used ambiguously in the visualization literature [Mun14]. Recent surveys tried to tackle this problem by identifying similarities across definitions and combining them in a more general statement [FIBK16] or by categorizing already existing definitions into more specific or general interpretations [BKC*13].

In summary, all glyph designs are graphical representations of data points, which can be positioned independently from each other. This flexibility in the layout is the biggest advantage of glyphs. They can be easily combined with other established visualizations opening space for various application areas. Geo-Spatial Data visualizations [AA04], Node-Link Diagrams [US09], TreeMaps [FFM12], or Scatter Plots [WG11] are just a few examples where glyph designs can enrich other visualization techniques with additional information about the data.

Although the design space of data glyphs is nearly endless [Mun14], some designs have received more research attention than others. Chernoff faces [Che73], star glyphs [SFGF72], or profiles [DTSS86] are prominent examples. However, in comparison to faces, star-like glyph designs and profiles are more often used in practice. Therefore, we want to focus on star-like glyphs and profiles to outline a very different approach to quality assessment and evaluation: *design recommendations*. For a better readability and didactic reasons, we enumerate recent and influential works on the evaluation aspect of design recommendations in the unified subsection *Summary of Evaluation approaches*.

Why Do We Need Quality Metrics for Glyphs?
In general, Star/Profile glyphs are closely related to parallel coordinate plots. They use a similar visual encoding to show dimensions and data values. Data lines are radiating from a center point to represent attribute dimensions. The length of each line is dependent on the
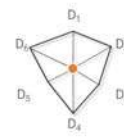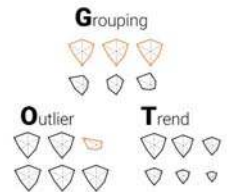

Figure 16: *Glyphs – Optimization Goals, Analysis Tasks & Visual Patterns*

underlying dimension value. The higher the data value, the longer the respective line. The endpoints of the data lines are connected to create a star-like shape. In comparison to parallel coordinates, the major differences are the reduced size, the circular layout of the axes and the presence of just a single data line in the plot.

As in Parallel Coordinates, the order of axis has a strong influence on the visual appearance of the individual stars and, therefore, need to be considered in the design process. Additionally, star glyphs can also be represented without the surrounding contour line [PG88], since this visual feature does not carry any information about the data. Adding color to the plot or highlighting certain visual features might also help to better solve the analysis task.

Profile glyphs are a more abstract term for small bar charts or line charts (i.e., sparklines [Tuf06]). They are easy to read and understand since they built upon a common mental model. Like in bar charts, the width of the bars, as well as their ordering, can be varied or single bars can be connected to show some trend information.

Although these two designs seem to be well-established, they still allow for some design variations. To come up with an optimal design is difficult, since the design of a glyph is a creative process with only limited guidance and nearly numberless design possibilities.

Typical Analysis Tasks for Glyphs
Data glyphs are used in various settings and for different analysis tasks. Based on Andrienko and Andrienko's task taxonomy [AA06], lookup tasks for single data values and similarity search are the most common analysis tasks followed by visual search and trend detection [FIBK16]. Therefore, the optimal glyph design strongly depends on the task at hand. Is it important to perceive the entire shape as a whole (like in synoptic tasks) or is the focus on reading individual visual features (like in elementary tasks).

Summary of Evaluation approaches
Glyph designs are a good example of visualization techniques, that strongly profit from design considerations based on results from quantitative user evaluation. Star glyphs profit from the following recommendations that can be used to guide the design process. The surrounding contour line should be removed from the design. Studies have shown that participants are more accurate when comparing the similarity between data points using stars without a contour line [FIB*14]. There are also guidelines for ordering the axes of stars. Results from experiments suggest avoiding salient shapes [KHW09]. This design consideration coincides with the clutter reduction quality metric proposed by Peng et al. [PWR04]. Additionally, the axes should be colored to reduce the negative influence from single spikes for visual classifications tasks [KHLW09].

To further improve the comparison between multiple stars, clustering results or statistical information should be added to the designs. Based on study results, researchers suggest adding the first and second principal component as additional axes to improve similarity comparisons [BS92]. Yang et al. [YPWR03] also proposed a quality metric to vary the angles between dimensions based on a hierarchical cluster analysis of the respective dimensions. Since no study has been conducted, this metric must be considered with caution.

However, the general public has to be careful about those recommendations, since all guidelines result from controlled experiments which are constructed to reflect specific conditions (e.g., analysis task, number of dimensions, layout). It is, therefore, difficult to generalize those findings [FIBK16].

Open Research Questions.
It would be interesting to transfer quality metrics from other visualization techniques to the data glyph domain. A good starting point can be Parallel Coordinates. Since Star Glyphs and Parallel Coordinates share many visual features, approaches for ordering dimensions could be adapted. Are Star Glyph specific orderings better compared to approaches used in Parallel Coordinate plots? Research has already made a first step in this direction by applying similar approaches to both visualization techniques [PWR04, YPWR03, HO12]. However, there is still much space for further research since the design space of data glyphs is huge.

7. Relational Data

In relational data, the units of observation are tuples rather than atomic entities. We here focus on the common case of binary relations represented as graphs. Graphs mark the crucial generalization from attribute data to relational data. A graph consists of a set of vertices and a set of edges which consist of ordered (in a directed graph) or unordered (in an undirected graph) pairs of vertices. Examples of data represented as graphs include dependency structures, hierarchies, and social networks.

Since graphs model relationships between pairs of entities, some relational data require even more general formalisms. Cleary, the entities linked by relationships and the relationships themselves can have additional attributes, thus increasing the complexity of the data. The composition of entities involved, the relationships they are subject to, and the attributes associated with either may change over time, giving rise to various classes of dynamic graphs.

Although graphs are often used synonymously with Node-Link Diagrams (7.1), there are many other graphical representations including Matrix Representations (7.2), inclusion drawings – of which TreeMaps (7.3) are a special case – contact representations (where edges are represented by touching vertex features), and visibility representations (where vertices are represented by sizeable features such as lines or boxes and the line-of-sight between adjacent pairs must not be blocked). Constrained variants such as orthogonal drawings, dendrograms, or arc-diagrams further add to the diversity.

7.1. Node-Link Diagrams

In Node-link diagrams, the vertices of a graph are represented as point-like features and its edges as line-like features connecting them. The main challenge in designing graph visualizations is that there are dependencies not only *between* different attributes, as in Multi- and High-dimensional Data, but also *within* variables

representing relationships. By fixing the layout to a grid, matrix representations avoid this difficulty but constrain relative locations. Treemaps, in contrast, apply only to the very restricted case of graphs that are connected and acyclic. Node-link diagrams, therefore, offer the most degrees of freedom, but also pose the most challenging layout problems [Tam13].
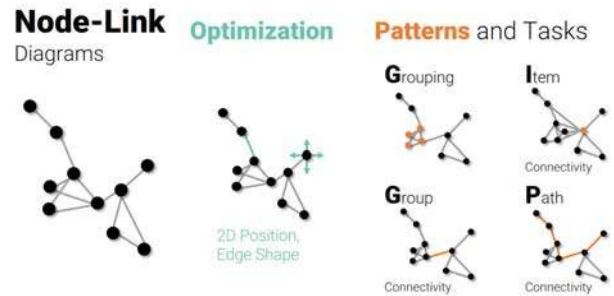

Figure 17: *Node-Link Diagrams – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Node-Link Diagrams?
A key feature in the visual understanding of graph drawings is the match between structural adjacency and graphical proximity. Common two-dimensional representations, however, often cannot depict the complex relationships faithfully [NEH13].

Because of the necessity to connect vertex features all these graphical representations face the problem of visual clutter. Since the layout is also contingent on the structure of relationships rather than just the values of attributes, issues such as occlusion, small angles between lines, or ambiguity of incidence arise. Many of these are present simultaneously in what is referred to as *hairball drawings*, i.e., drawings in which the structure of a (sub)graph is unrecognizable due to heavy overplotting.

The question, then, is what are the most important structural features to convey given a specific task, and where to compromise? Typical indicators for readability, traditionally referred to as aesthetics in the field of graph drawing, are the number of edge crossings, the angular resolution at vertices, edge complexity in terms of bends, the alignment of paths with straight lines connecting their origin and destination, the directional flow in directed graphs, and many more [DRSM15]. An indicator of the balancedness of a drawing as a whole is proposed in [AN02].

Typical Analysis Tasks for Node-Link Diagrams
Various tasks have been identified [LPP*06] and different designs have been proposed for different informational content (see, e.g., [BFW13] for a survey in the context of social networks).

The quintessential task of reading a graph is to determine whether two vertices are adjacent. More general versions of this task include reachability (the existence of a path between two vertices), structural distance, redundancy of connectivity, and the comparison of two vertices based on their relative position.

In addition to tasks parameterized with pairs of vertices, the task associated with single vertices include the size and structure of their neighborhood, their centrality, their membership in groups, and role as brokers between different regions of the graph.

Finally, structural features from the level of subgraphs to the

graph itself are studied. These include size and density, center-periphery divisions, variation in local cohesion (clustering, multi-centricity), hierarchies, symmetries, and substructures prevalence in graph mining.

The variety and complexity of these tasks increases further with the consideration of additional attributes and dynamics.

Summary of Approaches.
Graph layout algorithms are typically based on optimization of layout objectives that can be interpreted as quality criteria. The most widely used methods are referred to as force-directed algorithms and build on variations of the idea that adjacent vertices should be close to each other while non-adjacent vertices should be farther apart [Bra14]. As a consequence, these methods oftentimes also serve to display clustering and symmetries. While most of the objectives are formulated in terms of distances between vertices, a recent variant of stress majorization integrates dyad-wise direction objectives [WWS*18] and thus expands the expressiveness of such models.

A less common example is the introduction of a quality metric related to the visual group identification task [VBW17] in order to compare different layout adaptation strategies [NOB15].

Other graph layout algorithms are often designed to address quality metrics in isolation. Crossing, bend, and area minimization are prime examples. Many other techniques to reduce visual clutter in graph layouts such as edge bundling, confluent drawings, semantic substrates, or pivot graphs are based on alternative designs without an explicit quantitative assessment of layout quality.

Evaluation Methods for Node-Link Diagram Quality Metrics.
As stated above, quality metrics are incorporated directly or indirectly in optimization-based layout algorithms. For graphs conveying domain-specific information, it has been suggested to incorporate such information in terms of layout constraints, and restrict optimization to readability criteria [BKR*99]. Even then, however, the choices made are rarely informed by users studies.

Most often, readability metrics are evaluated in user studies on task performance subject to different graph visualizations [Pur97, HvW09, HIvWF11]. Comparative eye-tracking experiments are presented by Netzel et al. [NBW14] to study the effect of different link representations on the readability and by Burch et al. [BKH*11] to investigate different node-link layout strategies for their effectiveness. As a result, particular visualizations are considered more effective than others. An important concept to capture the difficulties users face is cognitive load [WPCM02, HEH09]. A range of works [BVKW11, HIvWF11] are trying to mitigate this problem and present their superiority over baseline methods in controlled user experiments. A range of other challenges, including cross-representation comparisons, arises when studying representations of graphs that change over time [AP16].

Open Research Questions.
The study of quality metrics in graph visualization has been mostly incidental and would benefit from more systematic approaches. Naturally, a principled and more comprehensive collection of quality metrics for graph visualizations would be helpful in the design of graph visualizations [DRSM15].

Quality-based evaluations so far have been concerned with ranking visualizations by effectiveness for the most part. More detailed

analysis of the various pairwise relationships between data, tasks, and users on the one side, and their individual and combined effects on quality metrics on the other side are yet to be devised. While efficiency is often used as an indicator of effectiveness, it could also be studied in its own right. Particularly for large graphs, the evaluation of metrics on samples may generalize to multiple problems [NHEM17].

Optimization need not involve the maximization of general readability, but could also tap more into human interpretative biases, for instance by deliberate misplacement in group detection tasks [M-BK96]. Conversely, how do metrics have to be adapted in order to correct, possibly in a personalized manner, for such biases? A thorough understanding of the impact of graph invariants on layout features relates not only to quality metrics, but would be highly beneficial for the controllability of visualization systems and the generation of instances for experimental user studies.

7.2. Matrix Representations
Matrix visualizations have an interesting property by design: They are able to show data patterns at a local and a global level-of-detail. This is facilitated by a simple, yet effective, layout algorithm in which data entities are shown in rows and columns and data attributes are mapped to the corresponding cell at the intersection of the corresponding row, respectively column.
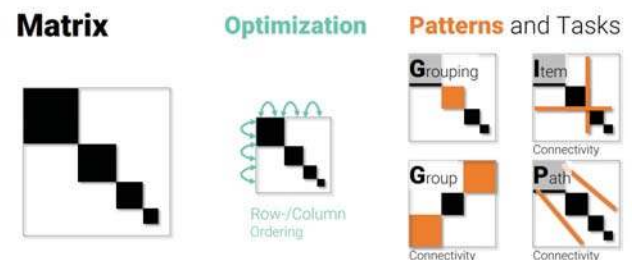

Figure 18: *Matrix – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Matrix Representations?
Matrix visualizations have been studied for centuries in many research and applications domains, such as archaeology, visualization research, and mathematics. However, Bertin [Ber73, Ber81] first recognized the *reorderable* matrix as an exploration tool, whose row-/column ordering reveals or potentially hides data insights. To understand *why* this is possible, it is important to note that the order of matrix rows and columns can be freely changed without changing the data in the matrix. Matrix reordering has been studied as an optimization problem, where the objective function is a heuristic to foster the visual appearance of visual base patterns depicted in Figure 19. In this paper, we will not focus on an all-embracing enumeration of matrix reordering algorithms. Interested readers can find a great variety of information on this topic in the recent matrix reordering survey of Behrisch et al. [BBR*16], Liiv's historic seriation overview [Lii10], or Wilkinson's and Friendly's "History of cluster heatmaps" [WF09].

Similar to the graph visualization community, there is no consensus of an objective function in the reordering literature. Hence, we cannot understand matrix reordering as a pure optimization problem,
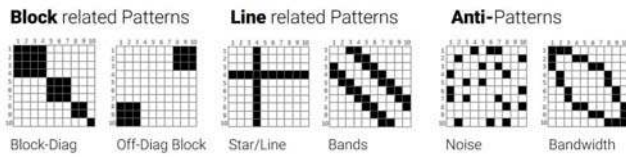
Figure 19: *Visual base patterns for matrix visualizations: Block-related patterns can be attributed to group and group relationship aspects, line-related patterns help to answer connectivity questions. If the matrix reordering algorithm is not able to reflect the underlying data topology or the algorithms' implementation characteristics dominate the data topology aspect, visual anti-patterns can emerge.*

but rather need to consider reordering algorithms according to the structures they reveal visually [BBR*16, Wil05].

Typical Analysis Tasks for Matrix Representations

Independent of an appropriate row-/column ordering, global network characteristics, such as size and density, or outstanding entities (e.g., highly connected entities), can be estimated [Ber81, GFC04]. Higher-level tasks, such as identifying groups, assessing distributions and cluster memberships, retrieving of (adjacency) relationships and general topology estimations, require a reordering of rows, respectively columns, to reveal higher order patterns [BBR*16]. Lee et al. [LPP*06] present a task taxonomy containing 12 low-level tasks and 13 high-level tasks for graph analysis and compare their applicability on five different visualization techniques.

Summary of Approaches.

The most challenging and focal question concerning matrix reordering is defining and evaluating which permutation of rows and columns is more effective than another. As depicted in Figure 20, more than one subjectively interesting reordering exists for a given matrix. Accordingly, the question arises which reordering result allows the user to perceive the data set's inner structure, patterns, regularity, and the overall trend the most.
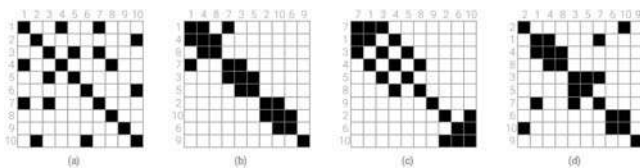

Figure 20: *Visual matrix reorderings (a) ordered randomly (b) and with three algorithms (c-e) revealing different patterns. Figure adapted from [BBR*16]*

Statistic global approaches to evaluate the visual quality of a matrix reordering exist: For example, *Column/row gradient* measures, such presented by Hubert et al. [Hub74], *Anti-Robinson Events/Deviations* [Che02, TLWC08], or the Inertia measure [CP05] relate all node-to-node distances (or edge weights) to their index-to-index distances. In that context, a matrix is said to have a *Robinsonian* form if the entries within each row and column are constantly decreasing the further away from the main diagonal. Especially,

blocks/clusters, diagonal structures can be promoted by optimizing for these objective functions [Rob51, Pet03]. Improving graph-theoretic metrics, such as the *Linear Arrangement* [Pet03, KH02] or the Hamilton path length [CP05] causes similar nodes to be placed close to one another and help to structure large matrices into smaller, homogeneous parts. Stress measures, such as presented by Niermann [Nie05], or the *Measure of Effectiveness* by McCormick [MDM-S69, MSW72] can be used to find locally dense areas, which are not necessarily centered around the main diagonal.

Wilkinson describes in "Grammar of Graphics" [Wil05] an entirely different approach inspired by calculating the retrieval performance. They are computing a Spearman correlation between row/column indices of an a-priori known visual base pattern and permuted indices from various matrix reordering algorithms. Behrisch et al. [BBH*17] present a pattern-driven quality metrics, called *Magnostics*, which generalizes Wilkinson's idea. In an image feature-driven approach averaged similarity scores to all base patterns (with varying noise levels) are computed. The resulting six-dimensional pattern-feature descriptor represents for each component the approximated visual similarity between the comparison matrix and all visual base patterns depicted in Figure 19.

Evaluation Methods for Matrix Quality Metrics.

The aforementioned quality metrics are mostly evaluated with comparative quantitative evaluation procedures. One generalizable example is Koren and Harel's presentation of the Multiscale matrix reordering algorithm [KH02] in which the processing time and the linear arrangement score are presented for selected competitors on the Petit testsuite [Pet03].

In case of the image feature-driven Magnostics approach, Behrisch et al. conducted an empirical evaluation consisting of four distinct sub-experiments (pattern response, -variability, -sensitivity and -discrimination) to validate that an engineered feature descriptor can be used to retrieve a specific base pattern.

Further influential for the understanding of matrix reorderings, and the visual results they produce, is the work of Mueller et al. [Mue04, MML07a, MML07b]. For example in [MML07a], a perception-inspired qualitative evaluation procedure was presented with the goal to measure (a) the *stability* of the visualizations for varying reorderings and (b) the *interpretability* or usefulness of the produced visual features. To a larger extent, Behrisch et al. [BBR*16] contribute to this discussion by focusing on the question which matrix reordering algorithm/-group is (algorithmically) able to depict a specific visual base pattern.

Open Research Questions.

The open research questions for this subfield are manifold. On the quality criterion side, all heuristics are focusing on global scores/aggregations. If a matrix is locally "well-organized", but in other parts rather noisy, global metrics will not be able to reflect the human's intuition. On the evaluation side, nearly no work has been devoted to proving that a human-engineered heuristic is able to mimic the human perception. Initial ideas to prove this congruency are presented in [LHT15, PKF*16] and should be adapted for this subdomain. Lastly, on the understanding side, more work needs to emphasize that visual distinct matrix reorderings of the same dataset. Matrix reorderings with pot. different pattern/task foci should be presented in a consistent format to the user.

7.3. TreeMaps

Treemaps are space-filling visualizations that make efficient use of the limited screen space to depict hierarchical data aspects. Shneiderman et al. [Shn92] were the first to develop a recursive subdivision schema for generating rectangular Treemaps. Each rectangle in a Treemap represents a node in a tree. Its area is designed to be proportional to the value of the node. To encode hierarchy, parent node rectangles enclose child rectangles [KHA10]. Generally, space-filling Treemap approaches allow one to remain comprehensible at much higher data densities than node-link diagrams.
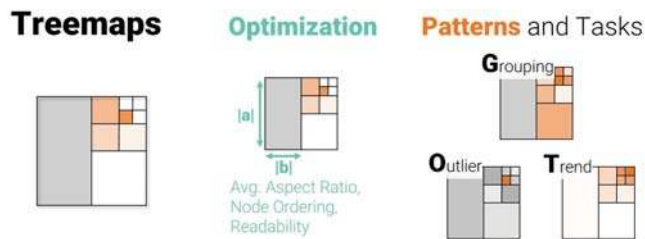

Figure 21: *Treemaps – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for TreeMaps?

Treemaps encode values using area. This basic construction paradigm is known to be less accurate than judgments of other visual encodings, such as length [Cle93a, CM84]. Various design alternatives have been proposed to overcome these issues, among others luminance of rectangles, 3D-effects with borders or different partitioning shapes. Yet, the creation of perceptually effective Treemaps requires carefully investigating the produced aspect ratio of the generated rectangles. This is also the most important quality criteria for Treemaps. Users find it difficult to compare regions with extreme aspect ratios [CHM82].

Typical Analysis Tasks for TreeMaps

Goldberg et al. [GH05] present a set of eight typical tasks for Treemap interfaces from the categories: Identification and counting, comparison using one or more criteria, advanced comparison and open-ended questions.

Summary of Approaches.

Bederson et al. [BSW02] present three quality criterion for Treemaps: Aspect Ratio, Change and Readability. The standard and accepted quality criterion is to minify the average aspect ratio (maximum of width/height and height/width) over all tree nodes. The lowest possible average aspect ratio is 1.0 representing a Treemap that only contains perfect squares.

The second and third quality criteria are more interesting, since they are inspired by perceptual workload considerations: The layout-distance-change quality criterion captures the idea that Treemaps should remain visually stable even if the underlying data is incrementally updated, such as often occurring in financial data analysis scenarios. Bederson defines the layout distance change function as the sum of all pairwise Euclidean distances between all rectangles defined by their 4-tuple (x, y, width, height). Consequently, a change of 0 would mean that no rectangles moved at all, and the more the rectangles are changed, the higher this metric will be. Tak and Cockburn [TC13], and Hahn et al. [HTMD14] presented alternative calculation procedures following the same quality criterion idea.
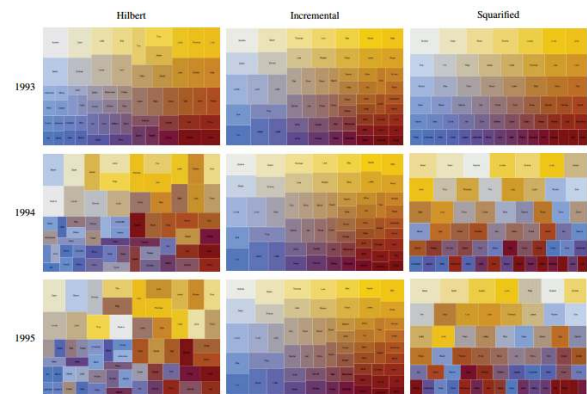

Figure 22: *The visual comparison is one of the prevailing evaluation methods for Treemaps, as shown here for Treemap stability approaches [SSV18]*

Finally, the readability criterion tries to capture how easily a user can to scan a layout to find a particular item. The measure tries to mimic the motion of the reader's eye fixation as the Treemap layout is scanned in order. To be precise, Bederson et al. consider the sequence of vectors needed to move along the centers of the layout rectangles in order, and count the number of angle changes between successive vectors that are greater than 0.1 radians (about 6 degrees) [BSW02, p. 9]. The resulting measure is equal to 1.0 in the most readable case, such as a slice-and-dice layout, and close to zero for a layout in which the order has been shuffled. Similarly, Tu and Shen [TS07] introduced the continuity metric which quantifies how often visual ordering does not match the data ordering in the enumeration of rectangles.

More recently, Ghoniem et al. [GCB*15] used a set of five quality criterion to evaluate their geography-preserving Treemap variant, amongst others average fragmentation (less is better), average angular displacement (less is better) or average distance displacement (less is better).

Evaluation Methods for TreeMaps Quality Metrics.

The Treemap layout algorithms are mostly compared with visual examples and an additional comparative quality score comparison based on various quality criteria.

A typical example is presented by Sondag in [SSV18] with the goal to show the (visual) superiority of their stable Treemap layout algorithm, as depicted in Figure 22.

Open Research Questions.

Bethge, Hahn, and Döllner recently presented an interesting dynamic Treemap layout algorithm which automatically combines eight existing Treemap layout algorithms [BHD17]. These novel hybrid layout algorithms are able to improve Treemap subregions based on data characteristics and quality notions. Even more interesting, in the same publication [BHD17], a neural network approach is applied to predict the layout quality metrics relative direction change, average distance change, and average aspect ratio.

Another open research question is the adaption of the existing quality criteria for non-rectangular shapes, such as Bubble-Treemaps [GSWD18] or Voronoi-Treemaps [BD05].

Figure 23: *Maps – Optimization Goals, Analysis Tasks & Visual Patterns*

8. Geo-Spatial Data

Geospatial datasets have a reference to our real world and geography [WGK10]. Many applications of geospatial data come from Earth Science and Physics. Other applications arise from the need to visualize census demographics, specific geospatial phenomena, and statistics, among other uses. Cartographers create maps of the real world to visualize and analyze spatial and non-spatial relationships among the data. The most popular visualizations for geospatial data are 2D maps, although there are other approaches such as 3D models and Earth globes [Ope18, GHD*17].

This work only surveys quality metrics for 2D maps because they are among the most popular geovis techniques. Due to the large scope of the techniques, we cover only an exemplary set of them, following the classification of [WGK10]: dot and density plots, line maps, area or choropleth maps, and cartograms [Tob04].

Why Do We Need Quality Metrics for Geo-Spatial Data Visualizations?

2D cartographic maps are the projected the representation of the Earth. The projection models used to create the 2D maps present a different kind of distortions that affect properties like shape, area, distance, and angles. Moreover, cartograms use distortion as part of the technique to express a given feature of a dataset. This distortion affects the data represented on the map directly or indirectly. There are references to previous work that show that distortion can also be used to mislead analysis tasks [Mon14]. Therefore the need for quality metrics. Another aspect that affects maps is clutter [RLMJ05, ED06b]. Previous efforts such as the PixelMaps [KPSN04] technique minimize overlapping of features and thus, clutter, and provide with performance metrics to measure its usefulness. In the case of choropleth maps, they need to take into account the perceptual limitations of color and the combination of other visual variables to prevent undesired color effects and artifacts that can affect the interpretation and analysis of patterns [War13].

Typical Analysis Tasks for GeoSpatial Data Visualizations

In contrast with other techniques, maps represent concepts with a semantic closer to humanity, such as cities, lands, roads, etc. For that reason analysis tasks are strongly attached to users and domains of application. Some tasks can be seen as common to several domains, for example: (1) map-reading of high-dimensional data, multi-resolution, with different LoDs. (2) Identify geographic "hot spots" where something interesting is happening. (3) Analyze statistical distributions over a geography in form of 2D fields (continuous) and grid (discrete) data. (4) Get insight about spatial patterns, clusters, systematic and random errors. (5) Comparison among multiple spatial objects or regions. Other common and generic tasks are to query a specific location, search for nearest neighbors, etc.

Summary of Approaches.

For general purpose maps, previous work focused on one of the most prominent characteristics of maps: distortion. All 2D map projections present distortion. The selection of the most effective map should be based on the quality criteria and usage of the map. The USGS [Sny82, KKESRIR00] provided a summary of map projections, properties, suitability, extent and location, and a qualitative measure of the amount of distortion each approach presents.

In the case of dot density maps and choropleth maps, the approaches tackle scalability, dynamic ranges issues, and cluttering. Pixelmaps [KPSN03, KPSN04] added a clustering pre-processing to make important data visible and provided with performance metrics for validation. Bertini et al. [BGS07] presented a novel density map technique based on statistical knowledge about the density distribution and a set of quality metrics that allows for validation of its effectiveness. Line maps cover a wide range of techniques such as flow maps, route maps, and networks. There is vast research done by Andrienko et al. [AAB*13] for the assessment of quality in trajectory visualization and movement data. Agrawala [Agr02] has done extensive work assessing the quality of route maps, just to mention a few works in the area.

Choropleth or Area maps use shades or colors to indicate areas, identify interesting hotspots and patterns. Brewer et al. [BMPH97] evaluated how the choice of a color scheme can influence the map interpretation, pattern analysis, and cluster identification, for a the particular case of the NCHS mortality atlas. Correll and Heer [CH17] presented a novel approach for the visualization of choropleth maps based on the Bayesian surprise technique. Their approach tackled misleading spatial patterns caused by sampling errors, artifacts or artificial patterns caused by the technique used to visualize the data, for example, normalization schemes such as percentages, per-capita rates, and z-scores. For cartograms, Tobler et al [Tob04] revisited definitions and algorithms used to construct them, as well as quality metrics to evaluate their performance and accuracy. Alam et al. [AKV15] surveyed how distortion can affect effectiveness. For a cartogram to be effective, it needs to be readable and recognizable with respect to their shape and neighbor areas. They proposed different classes of metrics: statistical distortion, topology distortion, shape distortion and complexity to assess quality. Nusrat et al. [NK16] surveyed three different types of quality metrics: topology accuracy, geography accuracy, and statistical accuracy of the geospatial data applied to Cartograms. They offered guides to mitigate inaccuracies:

  (i) Topological inaccuracies: by using simple interaction techniques, for example, "brushing" that highlights the neighbors of a selected state can help identify the correct topological relations.
 (ii) Geographic inaccuracies: by showing the cartogram alongside the undistorted geographical map with linking and brushing.
(iii) Statistical inaccuracies: by showing exact data values upon mouse-over events or listing data values next to the cartogram.

Wongsuphasawat [Won16] defined quality metrics for the layout design of Tile Grid Maps. Tong et al. [TRL*17] proposed cartograms or cartographic treemaps designed by combining space filling layouts and geocoordinates and proposed screen space filling and geospatial error metrics to validate their approach.

Figure 24: *Example of thematic maps. (a) Cartogram of the 2004 USA presidential election results, (b) Choropleth showing data from the 2,000 population census from USA, and (c) PixelMap showing telephone call volume analysis [SKNP04]*

Evaluation Methods for Geo-Spatial Data Visualization Quality Metrics.

There are strong efforts to quantify quality measures such as readability, effectiveness, expressiveness, performance, lookalikeness, uncertainty, to mention just a few. We have selected three main quality criteria: (1) accuracy (reliability of the map), (2) effectiveness (usefulness and effort required by the user to perform a task), and (3) expressiveness (data preservation; no data is added or lost).

Open Research Questions.

3D cartographic models or 3D world globes have been used for a long time but they became more popular lately with digitalization techniques and the introduction of interactive systems, like for example Google Earth, or 3D cartograms. Still, their use in analytical tasks continues being a hot debate in the visualization community in terms of how the navigation and interaction with the 3D map can help or slow down the cognitive process. Quality metrics could be very helpful in this endeavor. New augmented-reality, VR, and mixed-reality technologies bring a new dimension to the interactivity of maps, also resulting in a need for new quality metrics and perceptual studies to evaluate their effectiveness.

9. Sequential and Temporal Data

Sequential data is characterized by the serial order of data points in a sequence. The most prominent instance of sequential data is temporal and time-series data in which data involves consecutive time measures (e.g., economic developments or weather forecasts). Formally, temporal data can be defined as an ordered set of data points $d = f(t)$, each being a function of time. For a distinct timestamp t_i the corresponding data object can be represented as $d_i = f(t_i)$.

The analysis of time series data includes exploring temporal patterns, trend approximation and prediction, time-series segmentation, signal processing, and various other tasks. More concrete, MacEachren [Mac95] defines seven aspects of analysis concerning temporal data; namely, the existence of an entity, temporal location (when), a time interval (how long), temporal texture (how often), rate of change (how fast), sequence (in what order), and synchronization. There are two event types in a temporal analysis; namely discrete time points or time intervals (duration).

Furthermore, time can be structured into linear time (assumes starting point), cyclic time (periodic or recurring time events), and branching time (drifts like splits and merges of data points). To accommodate the diverse tasks related to sequential, especially temporal, data analysis, various visualization techniques have been proposed [AMM*07]. In this section, we are reviewing the two most common techniques with perceptual quality metrics.
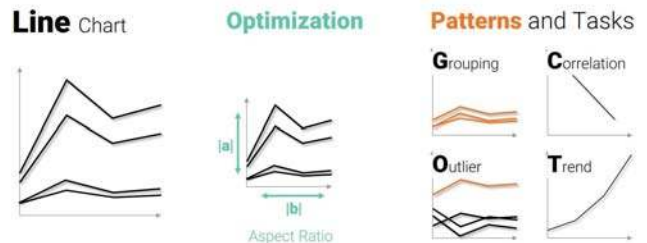

Figure 25: *Line Charts – Optimization Goals, Analysis Tasks & Visual Patterns*

9.1. Line Charts

A Line Chart is a sequential data visualization commonly used to depict temporal data, such as stock market prices, census data, or scientific data. Two axes in x- and y-direction are used for referencing the location of each data point in the coordinate system. Data points are shown as markers and connected by a straight line.

Why Do We Need Quality Metrics for Line Charts?

Line Charts deal with two primary design problems, that occur similarly in Scatter Plots (c.f., Section 6.1). First, when the number of time-series increases the typical superpositioning of lines becomes problematic. Second, choosing an inappropriate aspect ratio, i.e., height to width ratio, influences the orientations of the line segments, thus affecting the visual perception of trends or the accuracy of value judgments [Pal99, SED17].

Typical Analysis Tasks for Line Charts

Line Charts are used to analyze the temporal aspects of data. Accordingly, their related analysis tasks can be categorized into discrete comparison and trend assessment [ZT99]. More specifically, Saket et al. [SED17] mention (derived) value retrieval, filtering, finding extrema, sorting, distribution characterization, anomaly detection, finding clusters, and spotting of correlations as the typical analysis tasks for Line Charts. Whenever multiple time series are depicted, Javed et al. add slope differentiation and discrimination tasks to the list [JME10].

Summary of Approaches.

Interestingly, we found in our study that most of the heuristic QMs presented for Line Charts put their exclusive focus on the task of *trend analysis*. For many other tasks, Saket et al. [SED17] mentions a set of five statements guiding under which circumstances either Line Charts, Scatter Plots, Tables, Bar Charts, or Pie Charts should be preferred. The guidelines were derived from a crowdsourcing study with 203 participants measuring effectiveness and efficiency concerns. Regarding Line Charts the authors state that: *"[. . .] line charts and scatterplots have significantly higher accuracy and speed for finding correlations [. . .]"* Moreover, they state that for tasks that require readers to precisely identify the value of a specific data point Line Charts should be avoided.

On the heuristic QM side, Cleveland et al.'s [CMM88] *banking to 45°* is one of the fundamental quality criterion principles for improving the discriminability of adjacent line segments. Cleveland et al. [Cle93a, Cle93b, Cle94] proposed three quality metrics in the form of aspect ratio selection methods: median slope, average absolute orientation, and arc length weighted average absolute orientation.

Later, Guha and Cleveland [GC11] and Talbot et al. [TGH11] developed independently the resultant vector, respectively the arc length based methods, with the core intuition to minimize the arc length of the plotted curve while keeping the area of the plot constant. To foster visual perception Heer and Agrawala [HA06] proposed selecting the aspect ratio by maximizing the sum of squares of the angles between all pairs of segments in the plot. They developed two quality metrics, named the global orientation resolution method and the local orientation resolution. Additionally, Heer and Agrawala [HA06] present a frequency space analysis approach to decompose a time series into multiple components each of which can be banked independently. Recently, Wang et al. [WWZ*17] proposed a novel approach to combine existing aspect ratio selection methods, thus allowing to take advantage of their respective advantages, such as parameterization- or sampling invariances.

Evaluation Methods for Line Charts Quality Metrics.
Cleveland et al. [CMM88] were the first researchers who studied aspect ratio selection systematically. They conducted human-subject experiments and made the observation that the visual judgment of slope ratios between adjacent line segments is most accurate when the orientation resolution between segments is maximized. Since then, quality metrics for Line Charts are mostly presented in a *comparative (visual) inspection*; i.e., several data instances are depicted with several algorithm choices. This evaluation approach can be found in the most works in this subfield, e.g., Cleveland et al. [Cle93b], Heer and Agrawala [HA06], Talbot et al. [TGH11]

By using Han et al.'s [HWZ*16] problem formalization as line integrals, Wang et al. [WWZ*17] shows mathematically the similarity of several QM approaches: arc length based methods, average absolute orientation, and resultant vector. This kind of meta-analysis is important in a field that lacks standardized evaluation procedures.

As mentioned above, Saket et al. [SED17] present a comparative crowdsourcing study with 203 participants measuring effectiveness and efficiency of various Information Visualization techniques, including Line Charts, in ten distinct task scenarios (Find anomalies, Correlation, Distribution, Order, Filter, Find Clusters, Derived Value, Find Extrema, Retrieve Value, and Determine Range).

Open Research Questions.
Especially for Line Charts the first considerations for developing quality metrics stem from user study observations. Subsequently, more and more approaches have been developed and compared with only implicitly perceptual aspect considerations. Substantial future work remains to be done to close the gap between perceptual theory and the proposed practical methods for aspect ratio selection. On top of that, we see no perceptually-inspired QMs to guide aspect ratio selection for other tasks than trend analysis. Also interesting is the fact that the existing aspect ratio selection methods are designed for one curve, i.e., one time-series. Clutter-reduction is mainly achieved by faceting the problems into distinct, separate views [WGK10, 2nd Ed.; p. 293].

9.2. Stacked Graphs

Similar to a line chart, Stacked Graphs map sequential data onto a two-dimensional canvas. However, this type of visualization is tailored to show the temporal evolution of several data streams simultaneously. Each data sequence is mapped to an area segment.
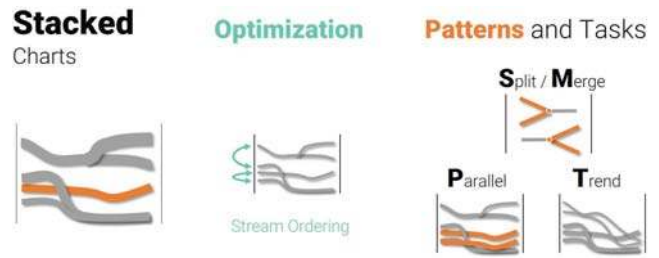

Figure 26: *Stacked Graphs – Optimization Goals, Analysis Tasks & Visual Patterns*

Multiple sequences are stacked on top of each other. Most commonly, stacked graphs are used to visualize temporal data, for example in the context of (news) topic evolution, where they are called ThemeRivers [HHWN02]. Stacked Graphs have two variations; on the one hand, plotting the values against a fixed straight axis – also referred to as Stacked Area Charts [Har00] – and, on the other hand, plotting the values against a varying central baseline – so-called *interior* Stacked Graphs [BW08].

Why Do We Need Quality Metrics for Stacked Graphs?
Havre et al. [HHWN02] introduced the ThemeRiver technique, an alternative to Stacked Bar Charts that enabled the tracking of the temporal evolution of attributes in a multivariate sequential dataset. Inspired by previous work, e.g., LifeLines [PMR*96], the main design rationale behind the proposed ThemeRiver visualization is to highlight the *temporal continuity* of every data value while showing the overall trend of the complete dataset (sum of the parts). To minimize the baseline shift for area segments, Havre et al. proposed a central (interior) baseline as an alternative to the bottom-aligned x-axis. This paper discussed two design challenges that influence the quality of the visualization; namely, the choice of color for every data segment (distinguishability), as well as, the calculation of faithful area portions through defining the right segment boundaries. Based on the metaphor of a thematic river, various visualization approaches have been proposed focusing on the analysis of the evolution of text corpora over time. Furthermore, as discussed by Byron and Wattenberg [BW08], in addition to the problems that occur for line charts, e.g., wrong aspect ratio, Stacked Graphs deal with two other issues that affect their visual perception. First, the order of the stacked area segments has a direct impact on the legibility of the visualization as is can hide patterns. Second, the shifting baseline is an integral aspect of Stacked Graphs which influences the comparability of data values and the detection of patterns. In addition, Cui et al. [CLT*11] discuss the need for optimizing the ordering of stacked segments to show splits and merges of the different layers.

Typical Analysis Tasks for Stacked Graphs
In contrast to Line Charts, Stacked Graphs are typically used to highlight the relation between multivariate data over time. If the temporal relation is linear, i.e., changing values over time, both approaches can be used to reveal patterns. However, to examine patterns within *branching* timelines, have proven to be more effective [Har00]. For example, Cui et al. [CLT*11] analyze temporal relations between segments in a Stream Graph, such as splits and merges, to follow topic evolutions over time. In addition, tasks like analyzing the relation of data values to others or to the overall graph are facilitated

by Stacked Graphs [HHWN02]. In addition to following splits and merges across layers, one central analysis tasks for Stacked Graphs is temporal trend detection [BW08].

Summary of Approaches.
Most Stacked Graphs are presented in the context of specific applications scenarios. Hence, the proposed approaches for enhancing their visual quality are tied to concrete algorithms that address the bulk of the required quality criteria at once. As one of the first approaches to address Stacked Graph optimization, Byron and Wattenberg [B-W08] describe layout algorithms based on the following steps; (1) computing the layer-geometry; (2) choosing the appropriate coloring scheme; (3) labeling of layers; and (4) layer ordering. One desirable characteristic for Stacked Graphs according to their design recommendation is to optimize the overall shape of the graph through avoiding "too flat or too spiky". They propose an implicit quality metric called a wiggle-minimizing method to avoid this issue. Extending their work for a specific domain problem, TIARA [LZP*09] was presented as an approach that segments a text corpus into topic layers evolving over time. This visualization is designed to satisfy three quality criteria to optimize the layer ordering; namely, "(1) minimizing layer distortion, (2) maximizing the available space within each layer to accommodate rich thematic content, and (3) ensuring visual proximity of layers to be proportional to their semantic similarity" [LZP*09]. The semantic similarity of two layers is computed by measuring the number of shared documents between their respective topics. As an alternative to the mirrored, central baseline, this approach proposed bottom-aligned graph to avoid "two artificial groups of topics" [LZP*09]. Addressing a similar analytical task, TextFlow [CLT*11] extends the previous visualization techniques to include splitting and merging concepts over time. This paper proposes a layout algorithm that is based on a three-level Directed Acyclic Graph to optimize the ordering of layers with the primary quality criterium to avoid edge crossings. This graph-based approach approximates different splits and merges using nodes that are ordered and placed according to their optimized position. Hereby, the notion of optimality is determined by the following criteria; reduce edge crossings; smooth topic layers (minimizing the edge length); and favor symmetry. Furthermore, Cui et al. [CLWW14] extended their previous work to support hierarchical data. They propose the usage of an incremental *evolutionary* tree cut algorithm in order to propagate the graph-based layout to different sub-layers, preserving the stability of the context in the visualization.

Evaluation Methods for Stacked Graphs Quality Metrics.
The quality of Stacked Graphs has been mostly determined by qualitative evaluation methods. As most Stacked Graph approaches were developed for the use case of topic evolution over time, the proposed evaluation methods are based on the users' perceived differences in the visualization quality. However, the optimization of the layer order, as well as, the graph baseline have been evaluated using defined measures. These are often used within the layout algorithms to determine the layer ordering and are, therefore, only useful as evaluation metrics in the broader context, i.e., when comparing the visual quality of different approaches. Byron and Wattenberg [BW08] describe quality criteria and quantitative metrics for optimizing the aesthetics and legibility of Stacked Graphs. They extend the baseline definition of Havre et al. [HHWN02], introducing a "weighted wiggle" function. Liu et al. [LZP*09] introduce the metric of "volatility" of a

topic layer to measure the fluctuation of a layer over time. This metric computes the standard deviation of the layer heights to determine a sorting strategy of the layers. For example, Liu et al. [LZP*09] propose sorting the least volatile layer close to the baseline. For splitting and merging layers over time, Cui et al. [CLT*11] propose to evaluate the number of edge crossings in order to determine the quality of the layout. Moreover, for hierarchical structures, Cui et al. [CLWW14] propose metrics based on the smoothness between adjacent tree cuts. They compute three values to determine the smoothness; namely, tree mapping (global tree cut energy function), normalized mutual information (similarity between adjacent tree cuts), and tree distance (avg. difference between two related cut nodes of the adjacent trees). In addition to their quantitative evaluation, this paper measures the quality of their proposed visualization based on the perceived stability of the generated graph. They performed user studies to measure layout similarity, visual clutter, and the support for topic tracking.

Open Research Questions.
As discussed by Byron and Wattenberg [BW08], there are various areas of open research with respect to measuring the visual quality of Stacked Graphs. Quantifying the quality of different layer reorderings, avoiding layer crossing, representing hierarchical information [CLWW14], and studying the effect of a shifting baseline on the graph legibility, are some of the open research challenges.

10. Text Data

The text is an inherently multivariate data source. Typically, when referring to textual data, we are regarding text corpora as a semi-structured source of information, molded by natural language processing and computational linguistics. Generally, linguistic and text visualizations [KK15, JFCS15] focus on revealing semantic information from the underlying raw text data. These can be the result of applying simple statistical processing (word-frequencies, bag-of-words models, etc.), more sophisticated linguistic processing (named-entity recognition, sentiment analysis, etc.), or combinations of both. Hence, most text visualization approaches rely on some text-mining approaches to transform the raw text into a structured data source [AdOP12, CC16].

These extracted data structures reveal different perspectives on the text data depending on the analysis tasks at hand. Based on the derived structure, various tailored techniques are utilized to visualize the data. For example, applying named-entity recognition can unveil the geospatial aspect of a text source through extracting mentioned locations, e.g., [CDW*16, GWS17]. Moreover, through linguistic computation, a text source can be transformed into a high-dimensional dataset of microlinguistic features, e.g., [KO07, ARLC*13]. In addition, parsing the relations of entities mentioned in a text yields a relational dataset or exposes hierarchical dependencies, e.g., [IM15, JEAGK17]. Furthermore, text corpora can be treated as sequential data, for example through using the timestamp of documents in a given collection (e.g., [EAGA*16, LYW*16]), or through parsing temporal information contained in a text document (e.g., [FBM16, EASG*17]).

All these potential transformations of text data make it a diverse data source. Through text-mining and other computations, the visualization of text data depends on the chosen extracted structure. In this section, we discuss two aspects of linguistic visualization

that are inherent to text data, namely legibility of Typographic Visualizations (10.1) and Tag Clouds (10.2). Readers may refer to the visualization types above to represent extracted time, geo, and high-dimensional features of text data.

10.1. Typographic Visualizations

Typography [Fel11] is the study of arranging type units of written language to display and render text. This arrangement determines the legibility and visual appeal of text, which is essential for various fields, such as, information visualization, design, and others. Typographic visualizations refer to all visualizations that rely on text sections, such as, labels in charts (e.g., [HWS17, EASG*17]) or typographic displays for close- and distant-reading (e.g., [GREA15, KJW*14]).

Figure 27: *Typographic Visualizations – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Typographic Visualizations?
Typography is used in different environments. For example, labeling a chart or a graph [BB16] has other constraints than displaying a complete text corpus in close- or distant-reading views [BGN08, BB14b]. To optimize the use of typography in visualizations, different parameters [SOK*16, ACS*17] have to be altered; these include typeface, font size, font weight, color, orientation, intensity (boldness), spacing, case, border, background, underline, shadow. The accurate setting of those parameters determines the legibility of text, thus, influences the understandability of visualizations. Furthermore, using the correct highlighting mechanisms facilitates the search and comparison in textual views or labels.

Typical Analysis Tasks for Typographic Visualizations
Optimizing typography is essential to perform three main tasks, namely: legibility (readability); localization and identification (highlight); and comparison (size). The first and most important task is the legibility (readability) of a written text, i.e., to understanding textual summaries [BB14b] or labels [BB16] in visualizations. Second, texts are rendered in visualizations to highlight certain aspects, therefore, localizing and identifying a specific text is an important task. For example, through using altering text and background color of keywords, visualization designers establish a visual linkage [SOK*16] between them. The last task is comparing text labels. Most commonly, labels are used to encode the importance of an attribute, for example, in their size. Setting the right typography for this task helps to avoid perceptual biases when comparing and interpreting text labels [ACS*17].

Summary of Approaches.
Similar to the design of Glyphs, the field of typography presents

quality considerations mainly by stating design recommendations and guidelines. Brath and Banissi [BB14b] comprehensively review different font attributes and their perceptual and analytic effect on visualizations. In their paper, they list a set of font-specific properties (typography dimensions) that can be adjusted depending on the analysis task. In addition, they discuss different usages of typographic attributes in visualizations to draw attention to words and regions of interest. They conclude that typographic attributes can be mapped back to visual channels [BB14a] to derive perceptual quality metrics and guidelines. As an example, font weight (corresponding to the size as a visual channel) would be more effective on ordered or quantitative data encodings than font family (corresponding to the shape as a visual channel). They also discuss integral (e.g., capitalization and italic) vs. separable (e.g., font weight and underline) typographic dimensions. In another work, Brath and Banissi [BB16] highlight the usage of optimized typography for the tasks of comparison and identification of labeled elements in set visualizations. To optimize the visual perception of text labels, they reviewed noticeable differences in the visual encoding of the typography. Another important task in the context of typography is typeface legibility. In her doctoral thesis [Bei09], Beier gives a comprehensive overview of the visibility and familiarity of typefaces and discusses the cognitive aspect of separation in visual perception. Her thesis examines the legibility of typefaces through extensive cognitive studies (also discussed in [BL13]) debunking the statement "readers read best what they read most" [Lic90].

Evaluation Methods for Typographic Visualization Quality Metrics.
Studying the perceptual effect of typography in information visualization has been a recent endeavor. So far, the typographic quality of text, visualizations, as well as, info-graphics, has been qualitatively evaluated through user studies. When it comes to the usage of typography in visualization, two notable user studies [SOK*16, ACS*17] evaluate common guidelines. In addition, Brath and Banissi [BB15] presented an information-theoretic approach quantifying the *lossiness* in visualizations based on evaluating the *fidelity* (number of unique levels perceivable for a particular visual attribute) in an encoding. Their approach measures the effectiveness of a particular set of visual encodings, which allows an efficient pruning of the design space of typographic dimensions. Moreover, the recent study by Strobelt et al. [SOK*16] empirically examined the effectiveness of highlighting techniques through a crowd-sourced user study, providing guidelines for nine different highlighting approaches. Their result shows the perceptual differences and superiority of the examined typographic dimensions and gives guidance on their utilization in information visualization. In a more recent work, Alexander et al. [ACS*17] study the perceptual biases in font size as a data encoding. As a commonly used typographic attribute, the font size has been utilized to encode data values in several visualization techniques. In their paper, they discuss results obtained across multiple experiments that consider factors such as the length and width of words and the height of the characters on the ability to judge comparative font sizes.

Open Research Questions.
Typography is an essential element of many visualizations and determines our understanding of the presented contexts. As such, studying the perceptual effects of the different typography dimensions

is a crucial task. Some preliminary approaches to measuring the quality of typographic visualizations for different tasks have been recently proposed. However, determining perceptually-superior typography variations for different visualization tasks is still an open research question. Another exciting research challenge is in optimizing typography for people with reading difficulties, e.g., dyslexia. Studies [RBY13] suggest that certain typefaces improve the accessibility (legibility) of textual data, which could be a significant factor for visualization design (similar to accessibility considerations for color blindness).

10.2. Tag Clouds

Tag (or Word) Clouds are a general-purpose visualization technique for text data with a broad usage in commercial tools and applications. This visualization technique has been proposed as an intuitive distant-reading technique that highlights a selection of prominent words or tags from a corpus to give an overview of the content [VW08]. This technique became widely popular around 2009, after the introduction of the *Wordle* web-platform [VWF09] for the automatic generation of word clouds from text snippets. As a visualization technique, word clouds have often been debated, not without controversy [SCH08, HR08], but nevertheless, have prevailed and continue to be optimized based on some quality criteria.

Figure 28: *Word/Tag Clouds – Optimization Goals, Analysis Tasks & Visual Patterns*

Why Do We Need Quality Metrics for Tag Clouds?
Tag Clouds map a selection of keywords onto a two-dimensional canvas. Their four principal visual variables are tag size (typically representing a numerical attribute, e.g., the keyword frequency); tag position (often algorithm-determined, in some cases semantically-anchored); tag color (often random, in some cases representing a categorical attribute); and tag orientation (mostly algorithm-determined). To enhance the perceptual quality of Tag Clouds, it is desirable to optimize the four visual variables, e.g., by determining optimal initial placements of keyword anchors and optimizing the text readability through optimal keyword orientations [WCB*18]. Furthermore, for incrementally changing data (e.g., topics over time) or comparable text corpora (e.g., news about the same event), one desirable optimization goal is the stability of word placement to facilitate keyword localization [CWL*10].

Typical Analysis Tasks for Tag Clouds
The primary analysis tasks supported by this visualization technique are exploration, overview, and summarization of document collections. Hence, Tag Clouds have been used as a distant-reading technique to get an initial impression of the content of a text corpus [JFCS15], substituting traditional list-based views. Tag Clouds have also been used for labeling automatically generated topics, as

discussed by Smith et al. [SLPS*17] who studied different visual representations for topic understanding. Lastly, Tag Clouds support the task of keyword search and corpus navigation [RGMM07].

Summary of Approaches.
Seifert et al. [SKK*08] propose a family of layout algorithms for Tag Clouds to optimize the packing and compactness of keywords. In their paper, they propose measures for evaluating the task-dependent usability of Tag Clouds and compare a set of layouts. Strobelt et al. [SSS*12] propose an overlap removal algorithm that combines a scan-line based method with concentric sorting to optimize the packing of keywords in a Tag Cloud. Through removing overlaps while maintaining the proximity relationships between tags, they maximize the perceived quality of the visualization. Another approach for layout optimization is proposed by Cui et al. [CWL*10]. In their paper on dynamic word cloud visualization, they present a technique tailored to ensure semantic coherence, while preserving the spatial stability of word clouds over time. Using an information-theoretic approach, they assess the perceptual quality of Tag Clouds through different criteria, for example, ensuring that words that appear or disappear simultaneously over time are spatially grouped (co-occurrence criterion). Using a feature vector for every criterion, a projection-based layout determines the initial placement of keywords, later, a force-directed layout is used to compute the packing of all tags. A more recent approach by Wang et al. [WCB*18] presented a stability-preserving technique for Tag Cloud layouts. Using a rigid body dynamics-based layout [Wit97], they introduce a set of constraints based on a selection of forces that define the quality of the Tag Clouds. These forces optimize, for example, neighborhood-preservation (stability), tag-centrality, or keyword-packing. The presented approach is tailored to retaining the stability of the spatial position of the keyword while editing a word cloud. In addition, semantic-preserving approaches are proposed to optimize the placement of keywords in Tag Clouds with respect to content clusters [WPW*11], geo-spatial information [BCL*16], or shape projections [PTT*12] .

Evaluation Methods for Tag Clouds Quality Metrics.
Similar to the evaluation of quality measures in typographic visualizations, the evaluation methods applied to measure quality in Tag Clouds are mostly based on qualitative user studies. Rivadeneira et al. [RGMM07] review the design space of Tag Clouds and present experimental evidence for evaluating their quality. In their studies, they considered optimizations based on typography (font weight, size, and color), as well as, word placement (sorting, clustering, spatial layout). They conclude that in designing Tag Clouds, the focus ought to be on optimizing the spatial layout (rather than other variables) as it demonstrates an effect on high-level cognitive processes, such as *impression formation* (overview). Another recommendation they give in their paper is to "consider the upper-left quadrant as a focal point" for placing tags that should be emphasized. Focusing specifically on the spatial layout, Lohmann et al. [LZT09] conducted a study to examine three common layout techniques for Tag Clouds on the user's perception and performance for different analysis tasks. Based on their findings, they recommend the following layout techniques; for the task of finding a specific tag: sequential layout with alphabetical sorting, for the task of finding the most popular tags: circular layout with decreasing popularity, and for the task of finding tags that belong to a certain topic: thematically

clustered layouts. Examining the effects of visual features, Bateman et al. [BGN08] measured perceptual effects on tag selection through varying nine visual properties of Tag Clouds, including typography and layout attributes. Their results indicate that particular attributes (e.g., font size, weight, and color) have a stronger influence on tag selection than others (e.g., number of pixels) that are "less likely to draw a viewer's attention." They also conclude that the centrality of keywords on the canvas impacts its prominence. In addition to layout and typography optimizations, Schrammel et al. [SLT09] evaluate the impact of semantically structured layouts on the search performance in Tag Clouds. Similarly, Halvey and Keane [HK07] examined the influence of varying visual attributes on search time. Furthermore, Jänicke and Chen [JC10] present a general quality metric based on visual salience. Their technique applies to all types of visualizations, as it measures how visual attention is directed in an image, favoring visualizations that direct the attention to the most relevant attributes. In their paper, they show how their technique is applied to measure the visual quality of Tag Clouds.

Open Research Questions.
One of the open challenges for Tag Cloud visualizations is the interpretation of spatial proximity [HMHS06]. Most quality criteria value the stability and compactness of a layout, however, often the produced visualizations form artifacts, such as perceiving two neighboring words as an n-gram (or sentence), setting the stage for false assumptions. Another aspect of quality that impacts Tag Clouds is the quality of the linguistic processing of the presented tags. To avoid repetitions of semantically-identical keywords in the visualizations, this visualization technique goes hand-in-hand with tailored linguistic processing methods. An opportunity for future research is in providing tighter integration of the visualization approaches and the linguistic computation, allowing for user-steerable refinements.

11. Opportunities and Future Directions

While this document describes existing approaches for the quality assessment of visualizations, there are still many opportunities to improve and extend existing metrics. In this Section, we report general findings and highlight promising future research directions.

11.1. General Findings and Discussion

In the following, we discuss common aspects of quality metrics that span across most visualization techniques.

Which QM favors which visual pattern? One of the central questions for QM design is how an effective instance of a particular visualization type should look like. This *understanding* is implicitly modeled into a heuristic algorithm trying to capture if the *subjective* QM designer's *expectation* of the visual structure is met. However, in exploratory analysis settings, it is unclear which QM to apply. Some QMs favor one visual pattern, others another. But, it remains to the user to guess which data or visual pattern is in the dataset. What is even worse is that a majority of QMs is presented or published for the purpose of quantitative algorithm evaluations *without* describing which visual pattern they prefer.

What are the extreme cases that a QM can deal with? And what happens if the specifications are not met? Only a few of our surveyed approaches have been systematically investigated for their noise (in-) variances and robustness toward skewed data distributions. However,

it is important that quality metrics can be applied independently of the quality of the data or the existence of patterns. A user should assume that no patterns exist in a dataset in case a quality metric does not provide a useful representation.

Is QM research transferable among visualization types? We found that some visualization subdomains share similar quality criterion. For example, many QMs developed for Scatter Plots and Line Charts (see: Section 6.1 and Section 9.1) are based on the *banking to 45 degree* principle. Another example are Scatter Plots and Parallel Coordinate Plots (see: Section 6.1 and Section 6.3) where the same clutter reduction techniques have been adapted for the respectively other visualization field (c.f., [ED06a, ED06b]). Further work should explore if research efforts can be transferred between visualization techniques and subdomains.

Are QMs equally descriptive? In the case that QMs for different visualization techniques are able to assess equally well the same visual patterns, then QMs could be used as visualization type recommenders. This, in turn, presumes that a standard set of base patterns has been accepted and established in the respective visualization subfield. This itself is a challenging question, since not only the type of dataset (e.g., hierarchical, high-dimensional, relational) influences the to-be-expected patterns but also the domain. For example, time-series matrices, such as used for financial data analysis [KNS*06] or high-throughput analysis matrices [LBK*18], are incomparable in terms of their visual pattern although they are matrix-like representations.

Evaluation of Quality Metrics. Notably, many works cited in this survey acknowledge and explicitly mention the fact that the evaluation of QMs is not backed up with perception-focused user studies. This statement holds explicitly for quantitative quality metrics. As mentioned in Section 3.3, design recommendations are mostly derived from qualitative and quantitative user studies.

We claim that both approaches are valid but eventually should be backed up with the respectively other approach. Heuristics should be evaluated for their perceptual aspects and proven to correspond to the humans' perceptual properties. This can be only done in structured large-scale user studies. Especially, crowdsourcing studies, such as in [HB10], allow for more and more (statistically) sound statements to be made. Design recommendations, in contrast, should be translated eventually into algorithms for deriving quantifiable heuristics. This step allows one to make design recommendations generally usable, comparable, and unambiguous.

Another important aspect for the evaluation of quality metrics is the availability of *perceptually-inspired* benchmark datasets. As one example, the Magnostics paper [BBH*17] presents a *synthetically* generated benchmark dataset comprising 5.570 matrix pattern images. For every of the five visual base patterns, and variations thereof, a gradual pattern degeneration schema is applied. This leads to perceptually increasingly vanishing/blurred instances of the visual base patterns. Each pattern degeneration function has the purpose to model quantifiable structural/topological-, algorithmic- or visual noise/artifacts into the pure base pattern images.

To address this issue, Schulz et al. [SNEA*16] propose generative data models for the validation, evaluation, and benchmark generation. In their paper, they survey various approaches that have been suggested to overcome the problem of the availability benchmark

datasets for different types of data. They argue for the use of generative gold-standard data for a standardized evaluation of visualization approaches, in particular, w.r.t.perceptual quality.

11.2. Open Research Questions and Promising Directions

Although Quality Metrics seem to be explored extensively, some fundamental questions cannot be answered thoroughly for most quality metrics:

Multi-Criterion QM. The current design of quality metrics follows *one* straight line. No bends or junctions are intended. However, our visualizations almost never expose just one pure visual pattern but present rather a mixture of several patterns. Accordingly, it is challenging to say under which circumstances QMs work and fail. What is needed are "flexible" QMs that adapt to the underlying dataset at hand and promote the main visual pattern, and, what is even more important, these *Multi-Criterion QM* should notify the user upon usage that their dataset contains more than just the main pattern and offer a faceted visual pattern space view. To give an example for Matrix-based or Graph visualization: A Multi-Criterion QM would first let the user see the primary expected pattern (e.g., data groupings) and then promote –optimally related– connectivity aspects (e.g., the connection between these groupings) to the analyst.

Task-Adapted QM. In interactive and exploratory systems, the notion of quality, especially in relation to the current analysis task, may change over time. However, the current QM approaches are not integrated into an exploration workflow and are not able to change their quality notion by adapting to the currently conducted task. To stick with our matrix-exploration example from above: An analyst might explore a large set of matrices with the goal to first filter out the potentially low number of matrices containing a visual pattern (recall is important), then secondly switches to the task to find data clusters, i.e., similarly behaving entities (recall and precision are equally important), but then switches tasks again to reason how these clusters related to each other (precision is of high importance). In such a scenario, not only the quality metric needs to be gradually exchanged during the exploration, but also importance for precision and recall must be adopted gradually.

Interactive and Human-Supported Quality Steering. Related to the two aforementioned aspects is interactive and human-supported quality assessment. Interactive and reactive systems should be able to facilitate the same exploration flexibility as the user in the process. Several noteworthy approaches have been presented in this young field with different foci on how quality metrics can be integrated in the exploration workflow:

Behrisch et al. [BKSS14] present a *relevance feedback* approach for a user-defined notion of interestingness in Scatter Plots. Users iteratively rank presented candidate views for their perceived interestingness. A gradually adapted classification model tries to mimic the current understanding of interestingness in a given feature space, while a so-called "*Decision Support System*" constantly monitors the user and assesses the relevance-driven search process for convergence and stability.

Another interesting approach is presented by Wongsuphasawat et al. [WMA*16] in their Voyager system. Voyager is designed as a mixed-initiative system, in which intelligent services and the user collaborate to achieve the analysis goals [Hor99] – an idea also

inherently incorporated in the Visual Analytics mantra [KAF*08]. Upon startup, the user is provided with a gallery of automatically generated visualizations for each (statistically) interesting data variable. The user navigates in the data space by a drill-down on one meaningful/expressive data variable and the underlying *Compass* recommendation engine enumerates, clusters, and ranks related visualizations according to both their underlying data properties and the resulting perceptual principles.

The interactive navigation and query definition for (complex) visual patterns is in the focus of the work of Shao et al. [SBS*14]. Confronted with a Scatter Plot pattern retrieval task the user draws a vague idea of an expected visual pattern into a canvas. Upon each stroke the system retrieves the most similar, respectively most dissimilar plots, an idea referred to as *guided-sketching*. Visually similar results are clustered and can be taken over to the canvas to adapt the search in this specific direction.

Machine Learning. Deep-learning based approaches have proven to be good visual pattern detectors. This could make *deep-learning based QM* a viable research direction. Two preconditions must hold: (1) a sufficiently large training dataset must be provided or generated, (2) an appropriate network structure has to be found that is able to deal not only with one expected visual pattern but rather a mixture-model of the pattern space. Recently, Sabour et al. [SFH17] have shown an interesting approach in which the activity vector of *groups of neurons*, so-called capsules, represents a specific type of entity, such as an object or an object part. This approach could be used for learning capsules, one for each visual pattern, and the network routing scheme decides which of the visual patterns are visually outstanding (have the most information content). While these approaches could lead to satisfactory results, proving their perceptual correspondence will be even harder since these approaches suffer inherently from the interpretability gap.

Closing the Gap to Higher-level Perceptual QM. As described before, mid-level perceptual quality metrics have the goal to retrieve interpretable visual patterns. In a successfully applied quality-metrics-driven Information Visualization pipeline, as depicted in Figure 1, the quality-metrics-driven automation part will help to reduce the analysts' cognitive overload by providing information about the dataset's inherent structure. While this is an active research field today, it makes sense to outline the benefits of developing multi-level perceptual and cognitive QMs. Multi-level quality metrics, e.g., combined of color perception QMs (low-level), pattern-detector QMs (mid-level), and memorability QMs (high-level) could find visually outstanding information with a high likelihood to be remembered. In another example, during the exploration of a relational dataset, a Multi-Level QM could decide to favor a graph representation over a matrix, since the analyst expressed more confidence in former analysis scenarios.

Meta-visualizations are inevitable for the users' *understanding and trust* in the quality metric assessment. These visualizations need to communicate the algorithmic certainty assessment, the percentage of the explored parameter space (especially, important in incremental scenarios), and the relationship between the result instances.

11.3. Limitations of this Survey

This work surveys mid-level perceptual quality metrics by motivating the needs and benefits of quality metrics in the respective

visualization subfield, summarizing the challenges and outlining analysis tasks supported by quality metrics in the literature. Our goal is to provide a central document where concepts from multiple visualization subdomains are enumerated and related, and their overarching concepts are discussed in contrast to each other.

While we discussed at length several alternatives to our present taxonomy, we finally opted to guide the reader through a structured questionnaire in each visualization section. We believe that the (missing) understanding of the visualization design challenges is a fundamental barrier to the effective use of visualizations in practice today. By providing a straightforward description of the problems and possible solutions in simple terms, we hope to help a wide audience better understand these algorithms and integrate them in future systems and libraries.

While we are trying to educate the user in the selection of QMs for a respective visualization type, a systematic answer to the question *"Which QM is the best one for my circumstances?"* remains extremely challenging. We decided against attempting to describe this matching formally. In particular, we do not think this is possible without considering domain-dependent, data-dependent, and user-dependent aspects.

12. Conclusion

This survey presents quality metric approaches for the field of Information Visualization. We summarize the efforts from 14 distinct visualization techniques/subfields along the five data types Multi- and High-Dimensional-, Relational-, Geo-Spatial, Sequential- and Temporal-, and Text data.

We found that the major research developments in the field are increasingly abandoning the idea of pure clutter reduction approaches and focus on visual pattern retrieval. This in turn has significant implications for visualization techniques and visual analytics in the exploration process. Within an integrated QM-driven automation, as depicted in Figure 1), the user will be guided to the primary (visual) patterns within the data *and* will be presented with a birds-eye perspective allowing to assess the dataset-inherent importance of each pattern. Thus, not only clustering-, but also outlier-, correlation-, and trend analysis tasks can be accomplished more effective and more efficient.

One of our core contributions of this work is that we formalize, unify, and exemplify the major QM vocabulary. In future, we can expect that such a unified understanding will enable a more structured work on this problem.

By gathering the knowledge in a central document, we hope to inspire more research to develop novel quality metric measurement strategies, more externalized and quantifiable criteria proven to mimic the analysts perceptual system, as well as novel exploration approaches to harness the power of QMs.

Acknowledgments

References

[AA04] ANDRIENKO N., ANDRIENKO G.: Interactive Visual Tools to Explore Spatio-temporal Variation. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2004), AVI '04, ACM, pp. 417–420. doi:10.1145/989863.989940. 17

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal data*. Springer, 2006. 17

[AAB*13] ANDRIENKO G., ANDRIENKO N., BAK P., KEIM D., WROBEL S.: *Visual Analytics of Movement*. Springer Publishing Company, Incorporated, 2013. 9, 22

[ABK98] ANKERST M., BERCHTOLD S., KEIM D. A.: Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Proceedings IEEE Symposium on Information Visualization* (Oct 1998), pp. 52–60, 153. doi:10.1109/INFVIS.1998.729559. 9, 13, 14, 15, 16

[ACS*17] ALEXANDER E. C., CHANG C. C., SHIMABUKURO M., FRANCONERI S., COLLINS C., GLEICHER M.: Perceptual biases in font size as a data encoding. *IEEE Transactions on Visualization and Computer Graphics PP*, 99 (2017), 1–1. doi:10.1109/TVCG.2017.2723397. 9, 26

[AdOP12] ALENCAR A. B., DE OLIVEIRA M. C. F., PAULOVICH F. V.: Seeing beyond reading: a survey on visual text analytics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2*, 6 (2012), 476–492. doi:10.1002/widm.1071. 25

[AEL*09] ALBUQUERQUE G., EISEMANN M., LEHMANN D. J., THEISEL H., MAGNOR M. A.: Quality-Based Visualization Matrices. In *VMV* (2009), pp. 341–350. 9, 11, 12

[AEL*10] ALBUQUERQUE G., EISEMANN M., LEHMANN D. J., THEISEL H., MAGNOR M.: Improving the visual analysis of high-dimensional datasets using quality measures. In *2010 IEEE Symposium on Visual Analytics Science and Technology* (Oct 2010), pp. 19–26. doi:10.1109/VAST.2010.5652433. 6, 9, 15, 16

[Agr02] AGRAWALA M.: *Visualizing Route Maps*. PhD thesis, Stanford University, 2002. 9, 22

[AKV15] ALAM M. J., KOBOUROV S. G., VEERAMONI S.: Quantitative measures for cartogram generation techniques. *Computer Graphics Forum 34*, 3 (2015), 351–360. doi:10.1111/cgf.12647. 9, 22

[AMM*07] AIGNER W., MIKSCH S., MÜLLER W., SCHUMANN H., TOMINSKI C.: Visualizing time-oriented data—a systematic view. *Computers & Graphics 31*, 3 (2007), 401–409. doi:10.1016/j.cag.2007.01.030. 9, 23

[AN02] ADAMEC J., NEŠETŘIL J.: Towards an aesthetic invariant for graph drawing. In *Graph Drawing: 9th International Symposium, GD 2001 Vienna, Austria, September 23–26, 2001 Revised Papers* (2002), Mutzel P., Jünger M., Leipert S., (Eds.), Springer Berlin Heidelberg, pp. 287–296. doi:10.1007/3-540-45848-4_23. 9, 18

[And57] ANDERSON E.: A Semigraphical Method for the Analysis of Complex Problems. *Proceedings of the National Academy of Sciences of the United States of America 43*, 10 (1957), 923–927. 17

[AP16] ARCHAMBAULT D., PURCHASE H. C.: Can animation support the visualisation of dynamic graphs? *Information Sciences 330* (2016), 495–509. 19

[ARLC*13] ABDUL-RAHMAN A., LEIN J., COLES K., MAGUIRE E., MEYER M., WYNNE M., JOHNSON C. R., TREFETHEN A., CHEN M.: Rule-based visual mappings–with a case study on poetry visualization. In *Computer Graphics Forum* (2013), vol. 32, Wiley Online Library, pp. 381–390. 25

[AS16] AUPETIT M., SEDLMAIR M.: Sepme: 2002 new visual separation measures. In *2016 IEEE Pacific Visualization Symposium (PacificVis)* (April 2016), pp. 1–8. doi:10.1109/PACIFICVIS.2016.7465244. 9, 12

[BB14a] BRATH R., BANISSI E.: The design space of typeface. In *Proceedings of the 2014 IEEE Symposium on Information Visualization (VisWeek 2014)* (2014). 9, 26

[BB14b] BRATH R., BANISSI E.: Using font attributes in knowledge maps and information retrieval. In *Knowledge Maps and Information Retrieval workshop* (2014), pp. 23–30. 9, 26

[BB15] BRATH R., BANISSI E.: Evaluating lossiness and fidelity in information visualization. In *Visualization and Data Analysis 2015* (2015), vol. 9397. 9, 26

[BB16] BRATH R., BANISSI E.: Typographic sets: Labelled set elements with font attributes. In *SetVR@ Diagrams* (2016), pp. 29–43. 9, 26

[BBH*17] BEHRISCH M., BACH B., HUND M., DELZ M., VON RÜDEN L., FEKETE J.-D., SCHECK T.: Magnostics: Image-based Search of Interesting Matrix Views for Guided Network Exploration. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 31–40. doi:10.1109/TVCG.2016.2598467. 3, 9, 20, 28

[BBK*16] BORKIN M. A., BYLINSKII Z., KIM N. W., BAINBRIDGE C. M., YEH C. S., BORKIN D., PFISTER H., OLIVA A.: Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan 2016), 519–528. doi:10.1109/TVCG.2015.2467732. 6

[BBR*16] BEHRISCH M., BACH B., RICHE N. H., SCHRECK T., FEKETE J.-D.: Matrix Reordering Methods for Table and Network Visualization. *Computer Graphics Forum 35*, 3 (2016), 693–716. doi:10.1111/cgf.12935. 9, 19, 20

[BCL*16] BUCHIN K., CREEMERS D., LAZZAROTTO A., SPECKMANN B., WULMS J.: Geo word clouds. In *2016 IEEE Pacific Visualization Symposium (PacificVis)* (April 2016), pp. 144–151. doi:10.1109/PACIFICVIS.2016.7465262. 9, 27

[BD05] BALZER M., DEUSSEN O.: Voronoi treemaps. In *IEEE Symposium on Information Visualization (InfoVis 2005)* (2005), p. 7. doi:10.1109/INFOVIS.2005.40. 9, 21

[BDS05] BERTINI E., DELL'AQUILA L., SANTUCCI G.: Springview: Cooperation of radviz and parallel coordinates for view optimization and clutter reduction. In *Coordinated and Multiple Views in Exploratory Visualization, 2005.(CMV 2005). Proceedings. Third International Conference on* (2005), IEEE, pp. 22–29. 16

[Bei09] BEIER S.: *Typeface legibility: towards defining familiarity*. PhD thesis, Royal College of Art, 2009. 9, 26

[Bel61] BELLMAN R. E.: *Adaptive control processes: a guided tour*. Princeton university press, 1961. 9

[Ber73] BERTIN J.: *Sémiologie Graphique - Les diagrammes, les reseaux, les cartes*. Éditions Gauthier-Villars, 1973. 19

[Ber81] BERTIN J.: Théorie matricielle de la graphique. *Communication et langages 48*, 1 (1981), 62–74. doi:10.3406/colan.1981.1409. 1, 19, 20

[BFW13] BRANDES U., FREEMAN L. C., WAGNER D.: Social networks. In *Handbook of graph drawing visualization*, Tamassia R., (Ed.). CRC Press, 2013, pp. 805–839. 18

[BGETT99] BATTISTA G D., EADES P., TOLLIS I. G., TAMASSIA R.: *Graph drawing: algorithms for the visualization of graphs*. 1999. 9

[BGN08] BATEMAN S., GUTWIN C., NACENTA M.: Seeing things in the clouds: the effect of visual features on tag cloud selections. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia* (2008), ACM, pp. 193–202. 9, 26, 28

[BGRS99] BEYER K. S., GOLDSTEIN J., RAMAKRISHNAN R., SHAFT U.: When Is "Nearest Neighbor" Meaningful? In *Proceedings of the 7th International Conference on Database Theory* (1999), ICDT '99, Springer-Verlag, pp. 217–235. 10

[BGS07] BERTINI E., GIROLAMO A. D., SANTUCCI G.: See What You Know: Analyzing Data Distribution to Improve Density Map Visualization. In *Eurographics/ IEEE-VGTC Symp. on Visualization* (2007), Ynnerman K. M. T. M. A., (Ed.), Eurographics Assoc. doi:10.2312/VisSym/EuroVis07/163-170. 9, 22

[BHD17] BETHGE J., HAHN S., DÖLLNER J.: Improving Layout Quality by Mixing Treemap-Layouts Based on Data-Change Characteristics. In *Vision, Modeling & Visualization* (2017), Hullin M., Klein R., Schultz T., Yao A., (Eds.), The Eurographics Association. doi:10.2312/vmv.20171261. 9, 21

[BKC*13] BORGO R., KEHRER J., CHUNG D. H., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN M.: Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. In *EG STAR* (2013), Eurographics, pp. 39–63. doi:10.2312/conf/EG2013/stars/039-063. 17

[BKH*11] BURCH M., KONEVTSOVA N., HEINRICH J., HÖFERLIN M., WEISKOPF D.: Evaluation of traditional, orthogonal, and radial tree diagrams by an eye tracking study. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2440–2448. doi:10.1109/TVCG.2011.193. 19

[BKR*99] BRANDES U., KENIS P., RAAB J., SCHNEIDER V., WAGNER D.: Explorations into the visualization of policy networks. *Journal of Theoretical Politics 11*, 1 (1999), 75–106. 19

[BKSS14] BEHRISCH M., KORKMAZ F., SHAO L., SCHRECK T.: Feedback-driven interactive exploration of large multidimensional data supported by visual classifier. In *Proc. IEEE Conference on Visual Analytics Science and Technology* (2014), pp. 43–52. Peer-reviewed full paper. doi:10.1109/VAST.2014.7042480. 29

[BL13] BEIER S., LARSON K.: How does typeface familiarity affect reading performance and reader preference? *Information Design Journal 20*, 1 (2013), 16–31. 9, 26

[BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. doi:10.1109/TVCG.2013.124. 3

[BMPH97] BREWER C. A., MACEACHREN A. M., PICKLE L. W., HERRMANN D.: Mapping mortality: Evaluating color schemes for choropleth maps. *Annals of the Association of American Geographers 87*, 3 (1997), 411–438. doi:10.1111/1467-8306.00061. 9, 22

[Bra97] BRATH R.: Metrics for effective information visualization. In *IEEE Symposium on Information Visualization* (1997), IEEE Computer Society, pp. 108–111. 7

[Bra14] BRANDES U.: *Force-Directed Graph Drawing*, 2nd ed. Springer, Boston, MA, 2014, pp. 1–6. doi:10.1007/978-3-642-27848-8_648-1. 19

[BS92] BORG I., STAUFENBIEL T.: Performance of snow flakes, suns, and factorial suns in the graphical representation of multivariate data. *Multivariate Behavioral Research 27*, 1 (1992), 43–55. doi:10.1207/s15327906mbr2701_4. 9, 18

[BS04] BERTINI E., SANTUCCI G.: Quality metrics for 2d scatterplot graphics: Automatically reducing visual clutter. In *Smart Graphics* (2004). 3, 5, 9, 10

[BS05] BERTINI E., SANTUCCI G.: Improving 2d scatterplots effectiveness through sampling, displacement, and user perception. In *Ninth International Conference on Information Visualisation (IV'05)* (2005), pp. 826–834. doi:10.1109/IV.2005.62. 9, 10

[BSM*15] BERNARD J., STEIGER M., MITTELSTÄDT S., THUM S., KEIM D. A., KOHLHAMMER J.: A survey and task-based quality assessment of static 2d colormaps. In *Visualization and Data Analysis 2015, San Francisco, CA, USA, February 9-11, 2015* (2015), p. 93970M. doi:10.1117/12.2079841. 7

[BSW02] BEDERSON B. B., SHNEIDERMAN B., WATTENBERG M.: Ordered and quantum treemaps: Making effective use of 2d space to display hierarchies. *ACM Trans. Graph. 21*, 4 (2002), 833–854. doi:10.1145/571647.571649. 9, 21

[BTK11] BERTINI E., TATU A., KEIM D.: Quality metrics in high-dimensional data visualization: An overview and systematization. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2203–2212. doi:10.1109/TVCG.2011.229. 2, 7, 9, 12

[BVKW11] BURCH M., VEHLOW C., KONEVTSOVA N., WEISKOPF D.: Evaluating partially drawn links for directed graph edges. In *Graph*

Drawing - 19th International Symposium, GD 2011, Eindhoven, The Netherlands, September 21-23, 2011, Revised Selected Papers (2011), pp. 226–237. `doi:10.1007/978-3-642-25878-7_22`. 19

[BW08] BYRON L., WATTENBERG M.: Stacked graphs – geometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (Nov. 2008), 1245–1252. `doi:10.1109/TVCG.2008.166`. 9, 24, 25

[CC16] CAO N., CUI W.: Overview of text visualization techniques. In *Introduction to Text Visualization*. Springer, 2016, pp. 11–40. 25

[CDW*16] CHO I., DOU W., WANG D. X., SAUDA E., RIBARSKY W.: Vairoma: A visual analytics system for making sense of places, times, and events in roman history. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 210–219. 25

[CH17] CORRELL M., HEER J.: Surprise! bayesian weighting for de-biasing thematic maps. *IEEE Transactions on Visualization and Computer Graphics* (2017). 9, 22

[Che73] CHERNOFF H.: The Use of Faces to Represent Points in K-Dimensional Space Graphically. *Journal of the American Statistical Association* (1973), 361–368. 17

[Che02] CHEN C. H.: Generalized association plots: information visualization via iteratively generated correlation matrices. *Statistica Sinica 12* (2002), 7–29. 9, 20

[CHM82] CLEVELAND W. S., HARRIS C. S., MCGILL R.: Judgments of circle sizes on statistical maps. *Journal of the American Statistical Association 77*, 379 (1982), 541–547. 21

[CJ10] CHEN M., JAENICKE H.: An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1206–1215. 7

[Cle93a] CLEVELAND W. S.: A model for studying display methods of statistical graphics. *Journal of Computational and Graphical Statistics 2*, 4 (1993), 323–343. 9, 10, 21, 23

[Cle93b] CLEVELAND W. S.: *Visualizing Data*. Hobart Press, 1993. 9, 23, 24

[Cle94] CLEVELAND W.: *The Elements of Graphing Data*. AT&T Bell Laboratories, 1994. 9, 23

[CLT*11] CUI W., LIU S., TAN L., SHI C., SONG Y., GAO Z., QU H., TONG X.: Textflow: Towards better understanding of evolving topics in text. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2412–2421. 9, 24, 25

[CLWW14] CUI W., LIU S., WU Z., WEI H.: How hierarchical topics evolve in large text corpora. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2281–2290. 9, 25

[CM84] CLEVELAND W. S., MCGILL R.: Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association 79*, 387 (1984), 531–554. 6, 21

[CMM88] CLEVELAND W. S., MCGILL M. E., MCGILL R.: The shape parameter of a two-variable graph. *Journal of the American Statistical Association 83*, 402 (1988), 289–300. 23, 24

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B.: *Readings in information visualization - using vision to think*. Academic Press, 1999. 1, 2, 4

[CP05] CARAUX G., PINLOCHE S.: Permutmatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics 21*, 7 (2005), 1280–1281. 9, 20

[CVW11] CLAESSEN J. H., VAN WIJK J. J.: Flexible linked axes for multivariate data visualization. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2310–2316. 9, 14

[CWB*14] CHEN M., WALTON S., BERGER K., THIYAGALINGAM J., DUFFY B., FANG H., HOLLOWAY C., TREFETHEN A. E.: Visual multiplexing. *Computer Graphics Forum 33*, 3 (2014), 241–250. `doi:10.1111/cgf.12380`. 7

[CWL*10] CUI W., WU Y., LIU S., WEI F., ZHOU M. X., QU H.: Context preserving dynamic word cloud visualization. In *IEEE Pacific Visualization Symp. (PacificVis)* (2010), IEEE, pp. 121–128. 9, 27

[CWRY06] CUI Q., WARD M. O., RUNDENSTEINER E. A., YANG J.: Measuring data abstraction quality in multiresolution visualizations. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 709–716. 9, 13, 14

[DAW13] DANG T. N., ANAND A., WILKINSON L.: Timeseer: Scagnostics for high-dimensional time series. *IEEE Transactions on Visualization and Computer Graphics 19*, 3 (2013), 470–483. `doi:10.1109/TVCG.2012.128`. 11, 12

[DB79] DAVIES D. L., BOULDIN D. W.: A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, 2 (1979), 224–227. 16

[DBETT94] DI BATTISTA G., EADES P., TAMASSIA R., TOLLIS I. G.: Algorithms for drawing graphs: an annotated bibliography. *Computational Geometry 4*, 5 (1994), 235–282. 7

[DCFMFM10] DI CARO L., FRIAS-MARTINEZ V., FRIAS-MARTINEZ E.: Analyzing the role of dimension arrangement for data visualization in radviz. *Advances in Knowledge Discovery and Data Mining* (2010), 125–132. 9, 16

[DK10] DASGUPTA A., KOSARA R.: Pargnostics: Screen-space metrics for parallel coordinates. *Visualization and Computer Graphics, IEEE Transactions on 16*, 6 (2010), 1017–1026. `doi:10.1109/TVCG.2010.184`. 3, 9, 12, 13, 14

[DKT17] DHEERU D., KARRA TANISKIDOU E.: UCI machine learning repository, 2017. 12

[DLR09] DRAPER G. M., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. *IEEE Transactions on Visualization and Computer Graphics 15*, 5 (2009), 759–776. 15

[DPS02] DIAZ J., PETIT J., SERNA M.: A survey of graph layout problems. *ACM Computing Surveys 34*, 3 (2002), 313–356. `doi:10.1145/568522.568523`. 7

[DRSM15] DUNNE C., ROSS S. I., SHNEIDERMAN B., MARTINO M.: Readability metric feedback for aiding node-link visualization designers. *IBM Journal of Research and Development 59*, 2/3 (2015), 14:1–14:16. 9, 18, 19

[DTSS86] DU TOIT S. H., STEYN A. G. W., STUMPF R. H.: *Graphical Exploratory Data Analysis*. Springer, 1986. 17

[DW14a] DANG T. N., WILKINSON L.: ScagExplorer: Exploring Scatterplots by Their Scagnostics. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE* (2014), pp. 73–80. `doi:10.1109/PacificVis.2014.42`. 11, 12

[DW14b] DANG T. N., WILKINSON L.: Transforming scagnostics to reveal hidden features. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1624–1632. 11, 12

[EAGA*16] EL-ASSADY M., GOLD V., ACEVEDO C., COLLINS C., KEIM D.: Contovi: Multi-party conversation exploration using topic-space views. In *Computer Graphics Forum* (2016), vol. 35 - 3, Wiley Online Library, pp. 431–440. 25

[EAM11] EISEMANN M., ALBUQUERQUE G., MAGNOR M.: Data driven color mapping. In *Proceedings of EuroVA: International Workshop on Visual Analytics, Bergen, Norway* (2011). 6, 7

[EASG*17] EL-ASSADY M., SEVASTJANOVA R., GIPP B., KEIM D., COLLINS C.: Nerex: Named-entity relationship exploration in multi-party conversations. *Computer Graphics Forum 36*, 3 (2017), 213–225. `doi:10.1111/cgf.13181`. 25, 26

[EBD05] ELLIS G., BERTINI E., DIX A.: The sampling lens: making sense of saturated visualisations. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems* (2005), ACM, pp. 1351–1354. 13

[ED06a] ELLIS G., DIX A.: Enabling automatic clutter reduction in parallel coordinate plots. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 717–724. 9, 13, 28

[ED06b] ELLIS G., DIX A.: The plot, the clutter, the sampling and its lens: Occlusion measures for automatic clutter reduction. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2006), AVI '06, ACM, pp. 266–269. doi:10.1145/1133265.1133318. 9, 13, 14, 22, 28

[ED07] ELLIS G., DIX A.: A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics* (2007). 2, 7, 9, 10, 12

[FBM16] FULDA J., BREHMEL M., MUNZNER T.: Timelinecurator: Interactive authoring of visual timelines from unstructured text. *IEEE Trans. on Visualization & Computer Graphics 22*, 1 (2016), 300–309. 25

[FBT*10] FERDOSI B. J., BUDDELMEIJER H., TRAGER S., WILKINSON M. H., ROERDINK J. B.: Finding and visualizing relevant subspaces for clustering high-dimensional astronomical data using connected morphological operators. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on* (2010), IEEE, pp. 35–42. 13

[Fel11] FELICI J.: *The complete manual of typography: a guide to setting perfect type*. Adobe Press, 2011. 26

[FFM12] FISCHER F., FUCHS J., MANSMANN F.: Clockmap: Enhancing circular treemaps with temporal glyphs for time-series data. *Proc. EuroVis Short Papers, Eurographics* (2012), 97–101. 17

[FFM*13] FUCHS J., FISCHER F., MANSMANN F., BERTINI E., ISENBERG P.: Evaluation of alternative glyph designs for time series data in a small multiple setting. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), ACM, pp. 3237–3246. 9

[FHSW13] FINK M., HAUNERT J.-H., SPOERHASE J., WOLFF A.: Selecting the aspect ratio of a scatter plot based on its delaunay triangulation. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2326–2335. doi:10.1109/TVCG.2013.187. 9, 10

[FIB*14] FUCHS J., ISENBERG P., BEZERIANOS A., FISCHER F., BERTINI E.: The influence of contour on similarity perception of star glyphs. *IEEE transactions on visualization and computer graphics 20*, 12 (2014), 2251–2260. 9, 17

[FIBK16] FUCHS J., ISENBERG P., BEZERIANOS A., KEIM D.: A systematic review of experimental studies on data glyphs. *IEEE Transactions on Visualization and Computer Graphics PP*, 99 (2016), 1–1. doi:10.1109/TVCG.2016.2549018. 17, 18

[FJ07] FORSELL C., JOHANSSON J.: Task-based evaluation of multirelational 3d and standard 2d parallel coordinates. In *Visualization and Data Analysis 2007* (2007), vol. 6495, International Society for Optics and Photonics, p. 64950C. 9, 14

[FPSS96] FAYYAD U., PIATETSKY-SHAPIRO G., SMYTH P.: From data mining to knowledge discovery in databases. *AI magazine 17*, 3 (1996), 37. 13

[FR11] FERDOSI B. J., ROERDINK J. B.: Visualizing High-Dimensional Structures by Dimension Ordering and Filtering using Subspace Analysis. In *Computer Graphics Forum* (2011), vol. 30, Wiley Online Library, pp. 1121–1130. doi:10.1111/j.1467-8659.2011.01961.x/full. 9, 13, 14

[FT74] FRIEDMAN J., TUKEY J.: A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers C-23*, 9 (1974), 881–890. doi:10.1109/T-C.1974.224051. 9, 11

[GC11] GUHA S., CLEVELAND W.: *Perceptual, mathematical, and statistical properties of judging functional dependence on visual displays*. Tech. rep., Technical report, Purdue University Dep. of Stat., 2011. 9, 24

[GCB*15] GHONIEM M., CORNIL M., BROEKSEMA B., STEFAS M., OTJACQUES B.: Weighted maps: treemap visualization of geolocated quantitative data. In *Visualization and Data Analysis 2015, San Francisco, CA, USA, February 9-11, 2015* (2015), p. 93970G. doi:10.1117/12.2079420. 9, 21

[GFC04] GHONIEM M., FEKETE J.-D., CASTAGLIOLA P.: A comparison of the readability of graphs using node-link and matrix-based representations. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on* (2004), Ieee, pp. 17–24. 9, 20

[GH05] GOLDBERG J. H., HELFMAN J. I.: Enterprise network monitoring using treemaps. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2005), vol. 49 - 5, SAGE Publications Sage CA: Los Angeles, CA, pp. 671–675. 21

[GHD*17] GORELICK N., HANCHER M., DIXON M., ILYUSHCHENKO S., THAU D., MOORE R.: Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* (2017). doi:10.1016/j.rse.2017.06.031. 22

[GLS17] GRAMAZIO C. C., LAIDLAW D. H., SCHLOSS K. B.: Colorgorical: creating discriminable and preferable color palettes for information visualization. *IEEE Transactions on Visualization and Computer Graphics* (2017). 7

[GREA15] GOLD V., ROHRDANTZ C., EL-ASSADY M.: Exploratory Text Analysis using Lexical Episode Plots. In *Eurographics Conference on Visualization (EuroVis) - Short Papers* (2015), Bertini E., Kennedy J., Puppo E., (Eds.), The Eurographics Association. doi:10.2312/eurovisshort.20151130. 26

[GSWD18] GÖRTLER J., SCHULZ C., WEISKOPF D., DEUSSEN O.: Bubble treemaps for uncertainty visualization. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 719–728. doi:10.1109/TVCG.2017.2743959. 9, 21

[GWS17] GODWIN A., WANG Y., STASKO J. T.: TypoTweet Maps: Characterizing Urban Areas through Typographic Social Media Visualization. In *EuroVis 2017 - Short Papers* (2017), Kozlikova B., Schreck T., Wischgoll T., (Eds.), The Eurographics Association. doi:10.2312/eurovisshort.20171128. 25

[HA06] HEER J., AGRAWALA M.: Multi-scale banking to 45 degrees. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 701–708. doi:10.1109/TVCG.2006.163. 9, 24

[HAK00] HINNEBURG A., AGGARWAL C. C., KEIM D. A.: What Is the Nearest Neighbor in High Dimensional Spaces? In *Proceedings of the 26th International Conference on Very Large Data Bases* (2000), VLDB '00, Morgan Kaufmann Publishers Inc., pp. 506–515. 10

[Har00] HARRIS R. L.: *Information graphics: A comprehensive illustrated reference*. Oxford University Press, 2000. 24

[HB10] HEER J., BOSTOCK M.: Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010), ACM, pp. 203–212. 28

[HE12] HEALEY C., ENNS J.: Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics 18*, 7 (2012), 1170–1188. 6

[HEH09] HUANG W., EADES P., HONG S.-H.: Measuring effectiveness of graph visualizations: A cognitive load perspective. *Information Visualization 8*, 3 (2009), 139–152. 9, 19

[HGM*97] HOFFMAN P., GRINSTEIN G., MARX K., GROSSE I., STANLEY E.: Dna visual and analytic data mining. In *Visualization'97., Proceedings* (1997), IEEE, pp. 437–441. 15, 16

[HHB08] HAHSLER M., HORNIK K., BUCHTA C.: Getting things in order: An introduction to the r package seriation. *Journal of Statistical Software 25*, 3 (2008), 1–34. 9

[HHWN02] HAVRE S., HETZLER E., WHITNEY P., NOWELL L.: Themeriver: visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics 8*, 1 (Jan 2002), 9–20. doi:10.1109/2945.981848. 9, 24, 25

[Hil91] HILBERT D.: Ueber die stetige abbilding einer line auf ein flächenstück. *Mathematische Annalen 38*, 3 (1891), 459–460. 14

[HIvWF11] HOLTEN D., ISENBERG P., VAN WIJK J. J., FEKETE J.: An extended evaluation of the readability of tapered, animated, and textured directed-edge representations in node-link graphs. In *IEEE Pacific Visualization Symposium, PacificVis 2011, Hong Kong, China, 1-4 March, 2011* (2011), pp. 195–202. doi:10.1109/PACIFICVIS.2011.5742390. 19

658

[HK07]   HALVEY M. J., KEANE M. T.: An assessment of tag presentation techniques. In *Proceedings of the 16th International Conference on World Wide Web* (New York, NY, USA, 2007), WWW '07, ACM, pp. 1313–1314. `doi:10.1145/1242572.1242826`. 9, 28

[HLKW12]   HEINRICH J., LUO Y., KIRKPATRICK A. E., WEISKOPF D.: Evaluation of a Bundling Technique for Parallel Coordinates. In *GRAPP & IVAPP 2012: Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications, Rome, Italy, 24-26 February, 2012* (2012), pp. 594–602. 9, 12, 14

[HMHS06]   HASSAN-MONTERO Y., HERRERO-SOLANA V.: Improving tag-clouds as visual information retrieval interfaces. In *International conference on multidisciplinary information sciences and technologies* (2006), Citeseer, pp. 25–28. 28

[HO12]   HURLEY C. B., OLDFORD R. W.: Pairwise display of high-dimensional information via eulerian tours and hamiltonian decompositions. *Jour. of Computational & Graphical Statistics* (2012). 9, 14, 18

[Hor99]   HORVITZ E.: Principles of mixed-initiative user interfaces. In *Proceeding of the CHI '99 Conference on Human Factors in Computing Systems: The CHI is the Limit, Pittsburgh, PA, USA, May 15-20, 1999.* (1999), pp. 159–166. `doi:10.1145/302979.303030`. 29

[HR08]   HEARST M. A., ROSNER D.: Tag clouds: Data analysis tool or social signaller? In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual* (2008), IEEE, pp. 160–160. 9, 27

[HRC15]   HARRISON L., REINECKE K., CHANG R.: Infographic aesthetics: Designing for the first impression. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2015), CHI '15, ACM, pp. 1187–1190. `doi:10.1145/2702123.2702545`. 6

[HTMD14]   HAHN S., TRÜMPER J., MORITZ D., DÖLLNER J.: Visualization of varying hierarchies by stable layout of voronoi treemaps. In *Information Visualization Theory and Applications (IVAPP), 2014 International Conference on* (2014), IEEE, pp. 50–58. 9, 21

[Hub74]   HUBERT L.: Some applications of graph theory and related non-metric techniques to problems of approximate seriation the case of symmetric proximity measures. *British Journal of Mathematical and Statistical Psychology 27*, 2 (1974), 133–153. `doi:10.1111/j.2044-8317.1974.tb00534.x`. 20

[HvW09]   HOLTEN D., VAN WIJK J. J.: A user study on visualizing directed edges in graphs. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009, Boston, MA, USA, April 4-9, 2009* (2009), pp. 2299–2308. `doi:10.1145/1518701.1519054`. 19

[HVW10]   HOLTEN D., VAN WIJK J. J.: Evaluation of cluster identification performance for different PCP variants. In *Computer Graphics Forum* (2010), vol. 29, Wiley Online Library, pp. 793–802. 9, 14

[HW13]   HEINRICH J., WEISKOPF D.: State of the art of parallel coordinates. In *Eurographics 2013 - State of the Art Reports, Girona, Spain, May 6-10, 2013* (2013), pp. 95–116. `doi:10.2312/conf/EG2013/stars/095-116`. 9, 13

[HWS17]   HU M., WONGSUPHASAWAT K., STASKO J.: Visualizing social media content with sentree. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (Jan 2017), 621–630. `doi:10.1109/TVCG.2016.2598590`. 26

[HWZ*16]   HAN F., WANG Y., ZHANG J., DEUSSEN O., CHEN B.: Mathematical foundations of arc length-based aspect ratio selection. In *2016 IEEE Pacific Visualization Symposium (PacificVis)* (2016), IEEE, pp. 9–15. `doi:10.1109/PACIFICVIS.2016.7465245`. 9, 24

[IM15]   INGRAM S., MUNZNER T.: Dimensionality reduction for documents with nearest neighbor queries. *Neurocomputing 150* (2015), 557–569. 25

[Ins85]   INSELBERG A.: The plane with parallel coordinates. *The visual computer 1*, 2 (1985), 69–91. 12

[Ins09]   INSELBERG A.: *Parallel Coordinates*. Springer US, Boston, MA, 2009. `doi:10.1007/978-0-387-39940-9_262`. 12

[JC08]   JOHANSSON J., COOPER M.: A screen space quality method for data abstraction. In *Computer Graphics Forum* (2008), vol. 27, Wiley Online Library, pp. 1039–1046. 9, 13

[JC10]   JÄNICKE H., CHEN M.: A salience-based quality metric for visualization. In *Computer graphics forum* (2010), vol. 29 - 3, Wiley Online Library, pp. 1183–1192. `doi:10.1111/j.1467-8659.2009.01667.x`. 7, 9, 28

[JEAGK17]   JENTNER W., EL-ASSADY M., GIPP B., KEIM D. A.: Feature alignment for the analysis of verbatim text transcripts. In *EuroVA 2017: EuroVis Workshop on Visual Analytics* (2017), pp. 13–18. 25

[JF16]   JOHANSSON J., FORSELL C.: Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 579–588. 9, 14

[JFC14]   JOHANSSON J., FORSELL C., COOPER M.: On the usability of three-dimensional display in parallel coordinates: Evaluating the efficiency of identifying two-dimensional relationships. *Information Visualization 13*, 1 (2014), 29–41. 9, 14

[JFCS15]   JÄNICKE S., FRANZINI G., CHEEMA M. F., SCHEUERMANN G.: On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *Eurographics Conference on Visualization (EuroVis) - STARs* (2015), Viola R. B. F. G. I., (Ed.), Eurographics Assoc. `doi:10.2312/eurovisstar.20151113`. 25, 27

[JFLC08]   JOHANSSON J., FORSELL C., LIND M., COOPER M.: Perceiving patterns in parallel coordinates: determining thresholds for identification of relationships. *Information Visualization 7*, 2 (2008), 152–162. 9, 14

[JJ09]   JOHANSSON S., JOHANSSON J.: Interactive dimensionality reduction through user-defined combinations of quality metrics. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 993–1000. 5, 9, 12, 13, 14

[JM03]   JANKUN-KELLY T. J., MA K.: Moiregraphs: Radial focus+context visualization and interaction for graphs with visual nodes. In *9th IEEE Symposium on Information Visualization (InfoVis 2003), 20-21 October 2003, Seattle, WA, USA* (2003), pp. 59–66. `doi:10.1109/INFVIS.2003.1249009`. 15

[JME10]   JAVED W., MCDONNEL B., ELMQVIST N.: Graphical perception of multiple time series. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 927–934. 23

[KAF*08]   KEIM D. A., ANDRIENKO G. L., FEKETE J., GÖRG C., KOHLHAMMER J., MELANÇON G.: Visual analytics: Definition, process, and challenges. In *Information Visualization - Human-Centered Issues and Perspectives*, Kerren A., Stasko J. T., Fekete J., North C., (Eds.), vol. 4950 of *Lecture Notes in Computer Science*. Springer, 2008, pp. 154–175. `doi:10.1007/978-3-540-70956-5_7`. 3, 8, 29

[Kan00]   KANDOGAN E.: Star coordinates: A multi-dimensional visualization technique with uniform treatment of dimensions. In *Proc. IEEE Information Visualization Symposium* (2000), vol. 650, p. 22. 15, 16

[Kei00]   KEIM D. A.: Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics 6*, 1 (2000), 59–78. 9, 14, 15

[KH02]   KOREN Y., HAREL D.: A multi-scale algorithm for the linear arrangement problem. In *Revised Papers from the 28th International Workshop on Graph-Theoretic Concepts in Computer Science* (2002), WG '02, Springer-Verlag, pp. 296–309. 9, 20

[KHA10]   KONG N., HEER J., AGRAWALA M.: Perceptual guidelines for creating rectangular treemaps. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 990–998. `doi:10.1109/TVCG.2010.186`. 9, 21

[KHLW09]   KLIPPEL A., HARDISTY F., LI R., WEAVER C.: Colour-enhanced star plot glyphs: Can salient shape characteristics be overcome? *Cartographica: The International Journal for Geographic Information and Geovisualization 44*, 3 (2009), 217–231. 9, 17

[KHW09] KLIPPEL A., HARDISTY F., WEAVER C.: Star plots: How shape characteristics influence classification tasks. *Cartography and Geographic Information Science 36*, 2 (2009), 149–163. doi:10.1559/152304009788188808. 9, 17

[KJW*14] KOCH S., JOHN M., WÖRNER M., MÜLLER A., ERTL T.: Varifocalreader: In-depth visual analysis of large text documents. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec 2014), 1723–1732. doi:10.1109/TVCG.2014.2346677. 26

[KK15] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)* (April 2015), pp. 117–121. doi:10.1109/PACIFICVIS.2015.7156366. 25

[KKESRIR00] KENNEDY M., KOPP S., ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (REDLANDS C.: *Understanding Map Projections: GIS by ESRI*. ArcGIS 8 concept guides. ESRI, 2000. 22

[KNS*06] KEIM D. A., NIETZSCHMANN T., SCHELWIES N., SCHNEIDEWIND J., SCHRECK T., ZIEGLER H.: A spectral visualization system for analyzing financial time series data. In *EuroVis06: Joint Eurographics - IEEE VGTC Symposium on Visualization* (2006), pp. 195–202. doi:10.2312/VisSym/EuroVis06/195-202. 28

[KO07] KEIM D. A., OELKE D.: Literature fingerprinting: A new method for visual literary analysis. In *IEEE Symposium on Visual Analytics Science and Technology* (2007), VAST '07, pp. 115–122. 25

[KPSN03] KEIM D. A., PANSE C., SIPS M., NORTH S. C.: Pixelmaps: A new visual data mining approach for analyzing large spatial data sets. In *Third IEEE International Conference on Data Mining* (2003), pp. 565–568. 9, 22

[KPSN04] KEIM D. A., PANSE C., SIPS M., NORTH S. C.: Pixel based visual data mining of geo-spatial data. *Computers & Graphics 28*, 3 (2004), 327–344. doi:10.1016/j.cag.2004.03.022. 9, 22

[KW78] KRUSKAL J. B., WISH M.: *Multidimensional scaling*, vol. 11. Sage, 1978. 11

[LAE*12] LEHMANN D. J., ALBUQUERQUE G., EISEMANN M., MAGNOR M., THEISEL H.: Selecting coherent and relevant plots in large scatterplot matrices. *Computer Graphics Forum 31*, 6 (2012), 1895–1908. doi:10.1111/j.1467-8659.2012.03069.x. 9, 11, 12

[LBK*18] LEKSCHAS F., BACH B., KERPEDJIEV P., GEHLENBORG N., PFISTER H.: Hipiler: Visual exploration of large genome interaction matrices with interactive small multiples. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 522–531. doi:10.1109/TVCG.2017.2745978. 28

[LHT15] LEHMANN D. J., HUNDT S., THEISEL H.: A study on quality metrics vs. human perception: Can visual measures help us to filter visualizations of interest? *it - Information Technology 57*, 1 (2015), 11–21. doi:10.1515/itit-2014-1070. 7, 10, 12, 20

[LHZ16] LU L. F., HUANG M. L., ZHANG J.: Two axes re-ordering methods in parallel coordinates plots. *Journal of Visual Languages & Computing 33* (2016), 3–12. 9, 13, 14

[Lic90] LICKO Z.: Typeface designs. *Émigré 15* (1990), 8–13. 26

[Lii10] LIIV I.: Seriation and matrix reordering methods: An historical overview. *Statistical Analysis and Data Mining 3*, 2 (2010), 70–91. doi:10.1002/sam.10071. 19

[LJC09] LIND M., JOHANSSON J., COOPER M.: Many-to-many relational parallel coordinates displays. In *Information Visualisation, 2009 13th International Conference* (2009), IEEE, pp. 25–31. 9, 14

[LKZ*15] LEHMANN D. J., KEMMLER F., ZHYHALAVA T., KIRSCHKE M., THEISEL H.: Visualnostics: Visual guidance pictograms for analyzing projections of high-dimensional data. *Computer Graphics Forum 34*, 3 (2015), 291–300. doi:10.1111/cgf.12641.

[LMVW10] LI J., MARTENS J.-B., VAN WIJK J. J.: Judging correlation from scatterplots and parallel coordinate plots. *Information Visualization 9*, 1 (2010), 13–30. 12

[LPP*06] LEE B., PLAISANT C., PARR C. S., FEKETE J.-D., HENRY N.: Task taxonomy for graph visualization. In *Proc. of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization* (New York, NY, USA, 2006), BELIV '06, ACM, ACM, pp. 1–5. doi:10.1145/1168149.1168168. 3, 9, 18, 20

[LRN88] LEE RODGERS J., NICEWANDER W. A.: Thirteen ways to look at the correlation coefficient. *The American Statistician 42*, 1 (1988), 59–66. 14

[LYW*16] LIU S., YIN J., WANG X., CUI W., CAO K., PEI J.: Online visual analytics of text streams. *IEEE Transactions on Visualization and Computer Graphics 22*, 11 (2016), 2451–2466. 25

[LZP*09] LIU S., ZHOU M. X., PAN S., QIAN W., CAI W., LIAN X.: Interactive, topic-based visual text summarization and analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management* (2009), ACM, pp. 543–552. 9, 25

[LZT09] LOHMANN S., ZIEGLER J., TETZLAFF L.: Comparison of tag cloud layouts: Task-related performance and visual exploration. In *IFIP Conf. Human-Computer Interaction* (2009), Springer, pp. 392–404. 9, 27

[Mac86] MACKINLAY J.: Automating the design of graphical presentations of relational information. *ACM Trans. Graph. 5*, 2 (Apr. 1986), 110–141. doi:10.1145/22949.22950. 6

[Mac95] MACEACHREN A.: How maps work: Representation, visualization, and design. 9, 23

[May00] MAYRING P.: Qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 1*, 2 (2000). doi:10.17169/fqs-1.2.1089. 8

[MBK96] MCGRATH C., BLYTHE J., KRACKHARDT D.: Seeing groups in graph layouts. *Connections 19*, 2 (1996), 22–29. 19

[MDMS69] MCCORMICK W. T., DEUTSCH S. B., MARTIN J. J., SCHWEITZER P. J.: *Identification of Data Structures and Relationships by Matrix Reordering Techniques*. Tech. rep., DTIC, 1969. 9, 20

[MHD*18] MATZEN L. E., HAASS M. J., DIVIS K. M., WANG Z., WILSON A. T.: Data visualization saliency model: A tool for evaluating abstract data visualizations. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (Jan 2018), 563–573. doi:10.1109/TVCG.2017.2743939. 7

[MHNW97] MILLER N., HETZLER B., NAKAMURA G., WHITNEY P.: The need for metrics in visual information analysis. In *ACM Workshop on New Paradigms in Information Visualization and Manipulation* (1997), ACM, pp. 24–28. 7

[MHS07] MACKINLAY J., HANRAHAN P., STOLTE C.: Show me: Automatic presentation for visual analysis. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (Nov 2007), 1137–1144. doi:10.1109/TVCG.2007.70594. 3, 6

[MJSK15] MITTELSTÄDT S., JÄCKLE D., STOFFEL F., KEIM D. A.: Colorcat: Guided design of colormaps for combined analysis tasks. In *Proc. of the Eurographics Conference on Visualization (EuroVis 2015: Short Papers)* (2015), vol. 2. 7

[MK15] MITTELSTÄDT S., KEIM D. A.: Efficient contrast effect compensation with personalized perception models. *Computer Graphics Forum 34*, 3 (2015), 211–220. doi:10.1111/cgf.12633. 7

[MLDZ07] MRAMOR M., LEBAN G., DEMŠAR J., ZUPAN B.: Visualization-based cancer microarray data classification analysis. *Bioinformatics 23*, 16 (2007), 2147–2154. 16

[MML07a] MUELLER C., MARTIN B., LUMSDAINE A.: A comparison of vertex ordering algorithms for large graph visualization. In *Visualization, 2007. APVIS'07. 2007 6th International Asia-Pacific Symposium on* (2007), IEEE, pp. 141–148. 9, 20

[MML07b] MUELLER C., MARTIN B., LUMSDAINE A.: Interpreting large visual similarity matrices. *2007 6th International Asia-Pacific Symposium on Visualization* (2007), 149–152. doi:10.1109/APVIS.2007.329290. 9, 20

[Mon14] MONMONIER M.: *How to Lie with Maps*. University of Chicago Press, 2014. 9, 22

[Mor66] MORTON G. M.: *A computer oriented geodetic data base and a new technique in file sequencing*. International Business Machines Company New York, 1966. 15

[MPOW17] MICALLEF L., PALMAS G., OULASVIRTA A., WEINKAUF T.: Towards perceptual optimization of the visual design of scatterplots. *IEEE Transactions on Visualization and Computer Graphics 23*, 6 (2017), 1588–1599. doi:10.1109/TVCG.2017.2674978. 9, 10

[MSW72] MCCORMICK W. T., SCHWEITZER P. J., WHITE T. W.: Problem decomposition and data reorganization by a clustering technique. *Operations Research 20*, 5 (1972), 993–1009. arXiv:http://dx.doi.org/10.1287/opre.20.5.993, doi:10.1287/opre.20.5.993. 9, 20

[MTL18] MATUTE J., TELEA A. C., LINSEN L.: Skeleton-based scagnostics. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 542–552. doi:10.1109/TVCG.2017.2744339. 9, 11

[Mue04] MUELLER C.: Sparse matrix reordering algorithms for cluster identification. *Machine Learning in Bioinformatics* (2004). 9, 20

[Mun14] MUNZNER T.: *Visualization Analysis and Design*. A.K. Peters visualization series. A K Peters, 2014. 8, 17

[NBW14] NETZEL R., BURCH M., WEISKOPF D.: Comparative eye tracking study on node-link visualizations of trajectories. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2221–2230. doi:10.1109/TVCG.2014.2346420. 19

[NEH13] NGUYEN Q., EADES P., HONG S. H.: On the faithfulness of graph visualizations. In *2013 IEEE Pacific Visualization Symposium (PacificVis)* (2013), pp. 209–216. doi:10.1109/PacificVis.2013.6596147. 18

[NGC01] NAGESH H., GOIL S., CHOUDHARY A.: Adaptive grids for clustering massive data sets. In *Proceedings of the 2001 SIAM International Conference on Data Mining* (2001), SIAM, pp. 1–17. 14

[NHEM17] NGUYEN Q. H., HONG S.-H., EADES P., MEIDIANA A.: Proxy graph: Visual quality metrics of big graph sampling. *IEEE Transactions on Visualization and Computer Graphics 23*, 6 (2017), 1600–1611. 9, 19

[Nie05] NIERMANN S.: Optimizing the ordering of tables with evolutionary computation. *The American Statistician 59*, 1 (2005). 9, 20

[NK16] NUSRAT S., KOBOUROV S.: The state of the art in cartograms. *Computer Graphics Forum 35*, 3 (2016), 619–642. doi:10.1111/cgf.12932. 9, 22

[NOB15] NOCAJ A., ORTMANN M., BRANDES U.: Untangling the hairballs of multi-centered, small-world online social media networks. *Journal of Graph Algorithms and Applications 19*, 2 (2015), 595–618. doi:10.7155/jgaa.00370. 19

[Nor06] NORTH C.: Toward measuring visualization insight. *IEEE Computer Graphics and Applications 26*, 3 (2006), 6–9. 1

[NRS97] NIEDERMEIER R., REINHARDT K., SANDERS P.: Towards optimal locality in mesh-indexings. In *International Symposium on Fundamentals of Computation Theory* (1997), Springer, pp. 364–375. 14

[NŠ11] NOVÁKOVÁ L., ŠTĚPÁNKOVÁ O.: Visualization of trends using radviz. *Journal of Intelligent Information Systems 37*, 3 (2011), 355. 16

[NVE*17] NETZEL R., VUONG J., ENGELKE U., O'DONOGHUE S., WEISKOPF D., HEINRICH J.: Comparative eye-tracking evaluation of scatterplots and parallel coordinates. *Visual Informatics 1*, 2 (2017), 118 – 131. doi:https://doi.org/10.1016/j.visinf.2017.11.001. 9, 12, 14

[Ope18] OPENSTREETMAP: Planet dump from https://planet.osm.org . https://www.openstreetmap.org, 2018. 22

[Pal99] PALMER S. E.: *Vision science: Photons to phenomenology*. MIT press, 1999. 9, 23

[PBO*14] PALMAS G., BACHYNSKYI M., OULASVIRTA A., SEIDEL H. P., WEINKAUF T.: An edge-bundling layout for interactive parallel coordinates. In *IEEE Pacific Visualization Symposium* (2014), IEEE, pp. 57–64. 9, 14

[Pet03] PETIT J.: Experiments on the minimum linear arrangement problem. *Journal of Experimental Algorithmics 8* (2003). doi:10.1145/996546.996554. 9, 20

[PG88] PICKETT R. M., GRINSTEIN G. G.: Iconographic displays for visualizing multidimensional data. In *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics* (1988), vol. 514, p. 519. 17

[PKF*16] PANDEY A. V., KRAUSE J., FELIX C., BOY J., BERTINI E.: Towards understanding human similarity perception in the analysis of large sets of scatter plots. In *Conference on Human Factors in Computing Systems CHI* (2016), pp. 3659–3669. doi:10.1145/2858036.2858155. 7, 9, 10, 12, 20

[PL17] PELTONEN J., LIN Z.: Parallel Coordinate Plots for Neighbor Retrieval. In *VISIGRAPP (3: IVAPP)* (2017), pp. 40–51. 9, 13

[PMR*96] PLAISANT C., MILASH B., ROSE A., WIDOFF S., SHNEIDERMAN B.: Lifelines: visualizing personal histories. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1996), ACM, pp. 221–227. 24

[PTT*12] PAULOVICH F. V., TOLEDO F. M. B., TELLES G. P., MINGHIM R., NONATO L. G.: Semantic wordification of document collections. *Computer Graphics Forum 31*, 3pt3 (June 2012), 1145–1153. doi:10.1111/j.1467-8659.2012.03107.x. 9, 27

[Pur97] PURCHASE H.: Which aesthetic has the greatest effect on human understanding? In *Graph Drawing* (Berlin, Heidelberg, 1997), DiBattista G., (Ed.), Springer Berlin Heidelberg, pp. 248–261. 9, 19

[PWR04] PENG W., WARD M. O., RUNDENSTEINER E. A.: Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on* (2004), IEEE, pp. 89–96. 9, 13, 14, 17, 18

[PWV*18] POLATSEK P., WALDNER M., VIOLA I., KAPEC P., BENESOVA W.: Exploring visual attention and saliency modeling for task-based visual analysis. *Computers & Graphics 72* (2018), 26–38. doi:10.1016/j.cag.2018.01.010. 7

[RBY13] RELLO L., BAEZA-YATES R.: Good fonts for dyslexia. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (2013), ACM, p. 14. 27

[RGMM07] RIVADENEIRA A. W., GRUEN D. M., MULLER M. J., MILLEN D. R.: Getting our head in the clouds: Toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2007), CHI '07, ACM, pp. 995–998. doi:10.1145/1240624.1240775. 9, 27

[RLMJ05] ROSENHOLTZ R., LI Y., MANSFIELD J., JIN Z.: Feature congestion: A measure of display clutter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2005), CHI '05, ACM, pp. 761–770. doi:10.1145/1054972.1055078. 9, 22

[Rob51] ROBINSON W. S.: A method for chronologically ordering archaeological deposits. *American antiquity 16*, 4 (1951), 293–301. 9, 20

[RP66] ROSENFELD A., PFALTZ J. L.: Sequential operations in digital picture processing. *Journal of the ACM 13*, 4 (1966), 471–494. 13

[RSRDS16] RUBIO-SÁNCHEZ M., RAYA L., DIAZ F., SANCHEZ A.: A comparative study between radviz and star coordinates. *IEEE Trans. on Visualization & Computer Graphics 22*, 1 (2016), 619–628. 16, 17

[RZH12] ROSENBAUM R., ZHI J., HAMANN B.: Progressive parallel coordinates. In *Visualization Symposium (PacificVis), 2012 IEEE Pacific* (2012), IEEE, pp. 25–32. 9, 14

[SA15] SEDLMAIR M., AUPETIT M.: Data-driven evaluation of visual quality measures. *Computer Graphics Forum 34*, 3 (2015), 201–210. doi:10.1111/cgf.12632. 7, 12

[SBLC17] SHER V., BEMIS K. G., LICCARDI I., CHEN M.: An empirical study on the reliability of perceiving correlation indices using scatterplots. *Computer Graphics Forum 36*, 3 (2017), 61–72. doi:10.1111/cgf.13168. 9, 12

[SBS*14] SHAO L., BEHRISCH M., SCHRECK T., V. LANDESBERGER T., SCHERER M., BREMM S., KEIM D.: Guided sketching for visual search and exploration in large scatter plot spaces. In *Proc. EuroVA International Workshop on Visual Analytics* (2014), pp. 19–23. Peer-reviewed short paper. doi:10.2312/eurova.20141140. 29

[SC94] STRAUSS A., CORBIN J.: Grounded theory methodology. *Handbook of qualitative research 17* (1994), 273–85. 8

[SCH08] SINCLAIR J., CARDEW-HALL M.: The folksonomy tag cloud: when is it useful? *Information Science Journal 34*, 1 (2008), 15–29. 9, 27

[SED17] SAKET B., ENDERT A., DEMIRALP Ç.: Data and task based effectiveness of basic visualizations. *CoRR abs/1709.08546* (2017). arXiv:1709.08546. 23, 24

[SES16] SAKET B., ENDERT A., STASKO J.: Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proc. of 6th Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (New York, NY, USA, 2016), BELIV '16, ACM, pp. 133–142. doi:10.1145/2993901.2993903. 6

[SFGF72] SIEGEL J., FARRELL E., GOLDWYN R., FRIEDMAN H.: The Surgical Implications of Physiologic Patterns in Myocardial Infarction Shock. *Surgery 72*, 1 (1972), 126–141. 17

[SFH17] SABOUR S., FROSST N., HINTON G. E.: Dynamic routing between capsules. *CoRR abs/1710.09829* (2017). arXiv:1710.09829. 29

[SG16] SZAFIR D., GLEICHER M.: Visualization-aware color design. In *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Posters* (2016), Eurographics Association, pp. 97–99. 7

[SG18] SARIKAYA A., GLEICHER M.: Scatterplots: Tasks, data, and designs. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (Jan 2018). doi:10.1109/TVCG.2017.2744184. 9, 10, 11

[SGM08] SHARKO J., GRINSTEIN G., MARX K. A.: Vectorized radviz and its application to multiple cluster datasets. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008). 17

[Shn92] SHNEIDERMAN B.: Tree visualization with tree-maps: 2-d space-filling approach. *ACM Trans. Graph. 11*, 1 (1992), 92–99. doi:10.1145/102377.115768. 9, 21

[Shn96] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages* (1996), VL '96, IEEE Computer Society, pp. 336–343. 3

[SKK*08] SEIFERT C., KUMP B., KIENREICH W., GRANITZER G., GRANITZER M.: On the beauty and usability of tag clouds. In *Information Visualisation, 2008. IV'08. 12th International Conference* (2008), IEEE, pp. 17–25. 9, 27

[SKNP04] SIPS M., KEIM D. A., NORTH S. C., PANSE C.: Visual data mining in large geospatial point sets. *IEEE CG&A 24* (09 2004), 36–44. doi:10.1109/MCG.2004.41. 9, 23

[SLPS*17] SMITH A., LEE T. Y., POURSABZI-SANGDEH F., BOYD-GRABER J., SEPPI K., ELMQVIST N., FINDLATER L.: Evaluating visual representations for topic understanding and their effects on manually generated labels. *Transactions of the Association for Computational Linguistics 5* (2017), 1–15. 9, 27

[SLT09] SCHRAMMEL J., LEITNER M., TSCHELIGI M.: Semantically structured tag clouds: an empirical evaluation of clustered presentation approaches. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), ACM, pp. 2037–2040. 9, 28

[SNEA*16] SCHULZ C., NOCAJ A., EL-ASSADY M., FREY S., HLAWATSCH M., HUND M., KARCH G., NETZEL R., SCHÄTZLE C., BUTT M., ET AL.: Generative data models for validation and evaluation of visualization techniques. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (2016), ACM, pp. 112–124. 28

[SNLH09] SIPS M., NEUBERT B., LEWIS J., HANRAHAN P.: Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum 28*, 3 (2009), 831–838. 9, 11, 12

[Sny82] SNYDER J. P.: *Map projections used by the U.S. Geological Survey Series*. Tech. rep., U.S. Geological Survey Series, 1982. 9, 22

[SOK*16] STROBELT H., OELKE D., KWON B. C., SCHRECK T., PFISTER H.: Guidelines for effective usage of text highlighting techniques. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 489–498. doi:10.1109/TVCG.2015.2467759. 9, 26

[SSB*16] SHAO L., SCHLEICHER T., BEHRISCH M., SCHRECK T., SIPIRAN I., KEIM D. A.: Guiding the exploration of scatter plot data using motif-based interest measures. *Journal of Visual Languages and Computing 36* (2016), 1–12. URL: https://doi.org/10.1016/j.jvlc.2016.07.003, doi:10.1016/j.jvlc.2016.07.003. 9, 11, 12

[SSES17] SHAO L., SILVA N., EGGELING E., SCHRECK T.: Visual exploration of large scatter plot matrices by pattern recommendation based on eye tracking. In *Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics* (New York, NY, USA, 2017), ESIDA '17, ACM, pp. 9–16. doi:10.1145/3038462.3038463. 12

[SSK06] SCHNEIDEWIND J., SIPS M., KEIM D. A.: Pixnostics: Towards measuring the value of visualization. In *2006 IEEE Symposium On Visual Analytics Science And Technology* (2006), IEEE, pp. 199–206. 3, 6, 9, 15

[SSS*12] STROBELT H., SPICKER M., STOFFEL A., KEIM D., DEUSSEN O.: Rolled-out wordles: A heuristic method for overlap removal of 2d data representatives. In *Computer Graphics Forum* (2012), vol. 31 - 3pt3, Wiley Online Library, pp. 1135–1144. 9, 27

[SSV18] SONDAG M., SPECKMANN B., VERBEEK K.: Stable treemaps via local moves. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 729–738. doi:10.1109/TVCG.2017.2745140. 9, 21

[SSZ*17] SACHA D., SEDLMAIR M., ZHANG L., LEE J. A., PELTONEN J., WEISKOPF D., NORTH S. C., KEIM D. A.: What you see is what you can change: Human-centered machine learning by interactive visualization. *Neurocomputing 268* (2017), 164–175. doi:10.1016/j.neucom.2017.01.105. 3

[STMT12] SEDLMAIR M., TATU A., MUNZNER T., TORY M.: A taxonomy of visual cluster separation factors. *Computer Graphics Forum 31*, 3pt4 (June 2012), 1335–1344. doi:10.1111/j.1467-8659.2012.03125.x. 9, 10

[TAE*09] TATU A., ALBUQUERQUE G., EISEMANN M., SCHNEIDEWIND J., THEISEL H., MAGNOR M., KEIM D.: Combining automated analysis and visualization techniques for effective exploration of high-dimensional data. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on* (2009), IEEE, IEEE, pp. 59–66. 5, 9, 11, 12, 13, 14, 16

[TAE*11] TATU A., ALBUQUERQUE G., EISEMANN M., BAK P., THEISEL H., MAGNOR M., KEIM D.: Automated analytical methods to support visual exploration of high-dimensional data. *IEEE Transactions on Visualization & Computer Graphics 17*, 5 (2011), 584–597. 9, 11, 12

[Tam13] TAMASSIA R. (Ed.): *Handbook of Graph Drawing and Visualization*. CRC Press, 2013. 18

[TAS04] TOMINSKI C., ABELLO J., SCHUMANN H.: Axes-based visualizations with radial layouts. In *Proceedings of the 2004 ACM symposium on Applied computing* (2004), ACM, pp. 1242–1247. 15

[TBB*10] TATU A., BAK P., BERTINI E., KEIM D., SCHNEIDEWIND J.: Visual quality metrics and human perception: An initial study on 2d projections of large multidimensional data. In *Proc. Int. Conf. on Advanced Visual Interfaces* (2010), AVI '10, ACM, pp. 49–56. doi:10.1145/1842993.1843002. 3, 9, 12

[TC13] TAK S., COCKBURN A.: Enhanced spatial stability with hilbert and moore treemaps. *IEEE Transactions on Visualization and Computer Graphics 19*, 1 (jan 2013), 141–148. doi:10.1109/TVCG.2012.108. 9, 21

662

[TGH11] TALBOT J., GERTH J., HANRAHAN P.: Arc length-based aspect ratio selection. *IEEE Transactions on Visualization & Computer Graphics 17*, 12 (2011), 2276–2282. `doi:10.1109/TVCG.2011.167`. 9, 24

[TGM83] TUFTE E. R., GRAVES-MORRIS P.: *The visual display of quantitative information*, vol. 2. Graphics Press, 1983. 1, 7

[TLWC08] TIEN Y., LEE Y., WU H., CHEN C.: Methods for simultaneously identifying coherent local clusters with smooth global patterns in gene expression profiles. *BMC Bioinformatics 9* (2008). `doi:10.1186/1471-2105-9-155`. 20

[Tob04] TOBLER W.: Thirty five years of computer cartograms. *Annals of the Association of American Geographers 94*, 1 (2004), 58–73. `doi:10.1111/j.1467-8306.2004.09401004.x`. 9, 22

[TRL*17] TONG C., ROBERTS R., LARAMEE R. S., BERRIDGE D., THAYER D.: Cartographic Treemaps for Visualization of Public Healthcare Data. In *Computer Graphics and Visual Computing (CGVC)* (2017), Wan T. R., Vidal F., (Eds.), The Eurographics Association. `doi:10.2312/cgvc.20171276`. 22

[TS07] TU Y., SHEN H.: Visualizing changes of hierarchical data using treemaps. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1286–1293. `doi:10.1109/TVCG.2007.70529`. 9, 21

[TT85] TUKEY J. W., TUKEY P. A.: Computer Graphics and Exploratory Data Analysis: An Introduction. In *Proc. the Sixth Annual Conference and Exposition: Computer Graphics '85, Vol. III, Technical Sessions* (1985), Nat. Computer Graphics Association, pp. 773–785. 11

[Tuf06] TUFTE E. R.: Beautiful evidence. *New York* (2006). 17

[US09] UNGER A., SCHUMANN H.: Visual Support for the Understanding of Simulation Processes. In *Visualization Symposium, 2009. PacificVis '09. IEEE Pacific* (2009), pp. 57–64. `doi:10.1109/PACIFICVIS.2009.4906838`. 17

[VBW17] VEHLOW C., BECK F., WEISKOPF D.: Visualizing group structures in graphs: A survey. *Computer Graphics Forum 36*, 6 (2017), 201–225. 19

[VC62] VC H. P.: Method and means for recognizing complex patterns, Dec. 18 1962. US Patent 3,069,654. 13

[VW08] VIÉGAS F. B., WATTENBERG M.: Timelines: Tag clouds and the case for vernacular visualization. *interactions 15*, 4 (July 2008), 49–52. `doi:10.1145/1374489.1374501`. 27

[VWF09] VIEGAS F. B., WATTENBERG M., FEINBERG J.: Participatory visualization with wordle. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (Nov 2009), 1137–1144. `doi:10.1109/TVCG.2009.171`. 27

[WAG05] WILKINSON L., ANAND A., GROSSMAN R. L.: Graph-theoretic scagnostics. In *IEEE Symposium on Information Visualization* (2005), vol. 5, IEEE Computer Society, p. 21. 3, 6, 9, 11, 12, 13

[War13] WARE C.: *Information Visualization: Perception for Design*. Information Visualization: Perception for Design. Morgan Kaufmann Publishers Inc., 2013. 6, 22

[Wat05] WATTENBERG M.: A note on space-filling visualizations and space-filling curves. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.* (2005), IEEE, pp. 181–186. 14

[WCB*18] WANG Y., CHU X., BAO C., ZHU L., DEUSSEN O., CHEN B., SEDLMAIR M.: Edwordle: Consistency-preserving word cloud editing. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 647–656. `doi:10.1109/`. 9, 27

[WEG87] WOLD S., ESBENSEN K., GELADI P.: Principal component analysis. *Chemometrics and Intelligent Laboratory Systems 2*, 1 (1987), 37–52. `doi:10.1016/0169-7439(87)80084-9`. 10, 11

[WF09] WILKINSON L., FRIENDLY M.: The history of the cluster heat map. *The American Statistician* (2009). 9, 19

[WFC*18] WANG Y., FENG K., CHU X., ZHANG J., FU C., SEDLMAIR M., YU X., CHEN B.: A perception-driven approach to supervised dimensionality reduction for visualization. *IEEE Transactions on Visualization and Computer Graphics 24*, 5 (2018), 1828–1840. `doi:10.1109/TVCG.2017.2701829`. 5, 10

[WG11] WARD M. O., GUO Z.: Visual exploration of time-series data with shape space projections. In *Computer Graphics Forum* (2011), vol. 30 - 3, Wiley Online Library, pp. 701–710. 17

[WGK10] WARD M., GRINSTEIN G., KEIM D.: *Interactive Data Visualization: Foundations, Techniques, and Applications*. A. K. Peters, Ltd., Natick, MA, USA, 2010. 8, 9, 22, 24

[WHZ*18] WANG Y., HAN F., ZHU L., DEUSSEN O., CHEN B.: Line graph or scatter plot? automatic selection of methods for visualizing trends in time series. *IEEE Transactions on Visualization & Computer Graphics 24*, 2 (2018), 1141–1154. `doi:10.1109/TVCG.2017.2653106`. 5

[Wil05] WILKINSON L.: *The Grammar of Graphics (Statistics and Computing)*. Springer-Verlag New York, Inc., 2005. 9, 20

[Wit97] WITKIN A.: Physically based modeling: principles and practice constrained dynamics. *Computer graphics* (1997). 9, 27

[WLP*13] WALKER R., LEGG P. A., POP S., GENG Z., LARAMEE R. S., ROBERTS J. C.: Force-directed parallel coordinates. In *Information Visualisation (IV), 2013 17th International Conference* (2013), IEEE, pp. 36–44. 9, 14

[WMA*16] WONGSUPHASAWAT K., MORITZ D., ANAND A., MACKINLAY J. D., HOWE B., HEER J.: Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 649–658. `doi:10.1109/TVCG.2015.2467191`. 3, 29

[Won16] WONGSUPHASAWAT K.: Quality metrics for tile grid maps. In *IEEE Conference on Information Visualization 2016* (2016). 9, 22

[WPCM02] WARE C., PURCHASE H., COLPOYS L., MCGILL M.: Cognitive measurements of graph aesthetics. *Information visualization 1*, 2 (2002), 103–110. 7, 9, 19

[WPW*11] WU Y., PROVAN T., WEI F., LIU S., MA K.-L.: Semantic-preserving word clouds by seam carving. In *Computer Graphics Forum* (2011), vol. 30 - 3, Wiley Online Library, pp. 741–750. 9, 27

[WQM*17] WONGSUPHASAWAT K., QU Z., MORITZ D., CHANG R., OUK F., ANAND A., MACKINLAY J. D., HOWE B., HEER J.: Voyager 2: Augmenting visual analysis with partial view specifications. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, May 06-11, 2017.* (2017), pp. 2648–2659. `doi:10.1145/3025453.3025768`. 3

[WW08] WILKINSON L., WILLS G.: Scagnostics distributions. *Journal of Computational & Graphical Statistics 17*, 2 (2008), 473–491. `arXiv:https://doi.org/10.1198/106186008X320465`, `doi:10.1198/106186008X320465`. 11, 12

[WWS*18] WANG Y., WANG Y., SUN Y., ZHU L., LU K., FU C., SEDLMAIR M., DEUSSEN O., CHEN B.: Revisiting stress majorization as a unified framework for interactive constrained graph visualization. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 489–499. `doi:10.1109/TVCG.2017.2745919`. 9, 19

[WWZ*17] WANG Y., WANG Z., ZHU L., ZHANG J., FU C.-W., CHENG Z., TU C., CHEN B.: Is there a robust technique for selecting aspect ratios in line charts? *IEEE Transactions on Visualization and Computer Graphics* (2017). 9, 24

[YPF00] YANG-PELÉZ J., FLOWERS W. C.: Information content measures of visual displays. In *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on* (2000), IEEE, pp. 99–103. 7

[YPWR03] YANG J., PENG W., WARD M. O., RUNDENSTEINER E. A.: Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *IEEE Symposium Information Visualization* (2003), IEEE, pp. 105–112. 9, 13, 14, 18

[YRWG13] YUAN X., REN D., WANG Z., GUO C.: Dimension projection matrix/tree: Interactive subspace visual exploration and analysis of high dimensional data. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2625–2633. `doi:10.1109/TVCG.2013.150`. 11

[ZT99] ZACKS J., TVERSKY B.: Bars and lines: A study of graphic communication. *Memory & Cognition 27*, 6 (1999), 1073–1079. 23