

The Good, the Bad, and the Ugly: A Theoretical Framework for the Assessment of Continuous Colormaps

Roxana Bujack*

Los Alamos National Laboratory

Terece L. Turton†

University of Texas at Austin

Francesca Samsel‡

University of Texas at Austin

Colin Ware§

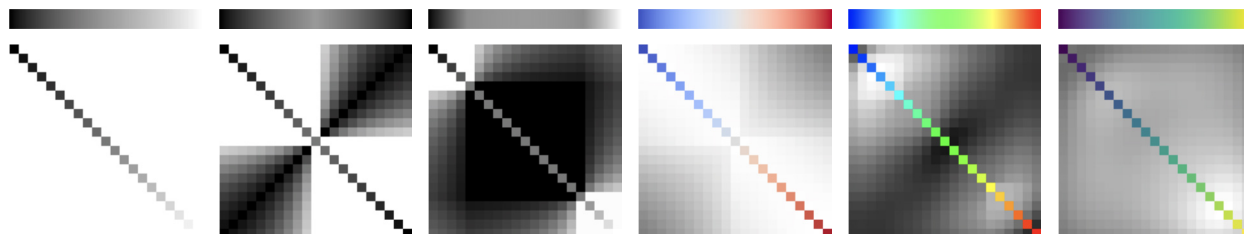
University of New Hampshire

David H. Rogers¶

Los Alamos National Laboratory

James Ahrens||

Los Alamos National Laboratory



(a) Greyscale. Constant speed between each pair of points. (b) Half greyscale. Constant speed in each half, drops to zero across the halves. (c) Flat greyscale. Constant grey in $[0.2, 0.8]$ results in a speed of zero. (d) Cool/warm divergent. Const. speed in each half, but not across halves. (e) Rainbow. High speed around blue, low around green. (f) Viridis. Low variations in speed result in its uniform appearance.

Fig. 1: Our six showcase colormaps and their corresponding matrices of the global speed. The luminance of each entry (i, k) represents how high the speed $V_{i,k}^{\Delta E^{76}} = \Delta E^{76}(x_i, x_k) / (t_k - t_i)$ is from the color $x(t_i)$ in the same row i to the color $x(t_k)$ in the same column k on the diagonal using the ΔE^{76} metric.

Abstract— A myriad of design rules for what constitutes a “good” colormap can be found in the literature. Some common rules include order, uniformity, and high discriminative power. However, the meaning of many of these terms is often ambiguous or open to interpretation. At times, different authors may use the same term to describe different concepts or the same rule is described by varying nomenclature. These ambiguities stand in the way of collaborative work, the design of experiments to assess the characteristics of colormaps, and automated colormap generation. In this paper, we review current and historical guidelines for colormap design. We propose a specified taxonomy and provide unambiguous mathematical definitions for the most common design rules.

Index Terms— colormap, survey, taxonomy, order, uniformity, discriminative power, smoothness, monotonicity, linearity, speed



1 MOTIVATION

Colormapping is a very old technique. Dating back to the 19th century, early colormaps were often based on prior experience, practical reasoning, or aesthetic opinion rather than an experimentally-based understanding of the human visual system and perception [18]. In our literature research, we found that even today, the assessment of good, bad, or ugly colormaps often relies on these values, partly because meaningful experiments can be difficult to design [81] or implement.

One possible reason for the difficulty in assessing and defining the characteristics that make an effective colormap is the lack of a common framework. Wainer and Francolini stressed the importance of a common language decades ago: “For if we do not have a vocabulary with which to discuss graphic concepts, we cannot discuss these concepts in an unambiguous manner.” Yet ambiguities continue.

At times, the same terms are used to describe different properties of a

color map. For example, *order* as used by Levkowitz and Herman [30] has a local and directed meaning: a color x_i should perceptually precede the next color. In contrast, the term *order* as used by Ware [83] has a global but undirected meaning. He requires that a user be able to sort colors picked from anywhere in the colormap, but does not distinguish whether they are sorted from low to high or vice versa. While *linearity* as used by Pizer [51] refers to the perceptual distance between neighboring colors, Levkowitz and Herman [30] use *linear* to require that the path of the colormap forms a straight line.

Alternately, different authors use different terms to refer to the same concept. Levkowitz and Herman [30] use the terminology *no boundaries*, Borland [8] uses *no sharp transitions*, while Moreland [46] uses the term *no Mach bands*. All of these seem to represent the same concept. Likewise, perceptual *uniformity* [30], *linearity* [52], *homogeneity* [82], *representative distances* [68], and *preservation of data relations* [41] are used interchangeably.

These ambiguities can impede experimental design, collaborative efforts, and attempts to automate the assessment, choice, or generation of colormaps. We hypothesize that the goal of creating a cross-discipline framework and language can best be achieved through that most fundamental of disciplines: mathematics. With this hypothesis, we agree with Resnikoff [54], who states “it has been generally recognized that a theory of color perception must be, both in form and content, a mathematical theory.” and Kindlmann and Scheidegger [26].

Collaborating with a mathematician, a physicist, an artist, a perceptual scientist, and a computer scientist, we strove to unify and identify nomenclature for the suggested rules in a way that minimizes interference with terminology in other fields. Additionally, we tried to find nomenclature that is expressive, intuitive, unambiguous, general,

*e-mail: bujack@lanl.gov

†e-mail: tlturton@cat.utexas.edu

‡e-mail: figs@cat.utexas.edu

§e-mail: cware@com.unh.edu

¶e-mail: dhr@lanl.gov

||e-mail: ahrens@lanl.gov

Manuscript received 31 Mar. 2017; accepted 1 Aug. 2017.

Date of publication 28 Aug. 2017; date of current version 1 Oct. 2017.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TVCG.2017.2743978

simple, and in accordance with the names that have been used in the literature to date. After long discussions, we were able to agree on the taxonomy used in this paper. The definitions are general in the sense that they are operational and independent from the colorspace, task, or data under consideration. The contributions of this paper are as follows:

- We start by cataloging the many suggested colormap design rules from the wide body of literature available.
- We categorize them into perceptual, mathematical, and operational rules.
- We interpret the perceptual rules based on how the authors used them and, where possible, we delineate this interpretation by means of a mathematical formula.
- We unify, clarify, and standardize the taxonomy by distinguishing different uses of the same terms and summarizing different terms for the same concept.
- Whenever possible, we associate perceptual concepts with mathematical concepts and suggest measures to evaluate them.
- We provide an online tool that allows the user to assess a colormap based on these measures.

Please note that the following items are specifically not attempted contributions of this paper:

- **We do not** suggest new rules to evaluate colormaps; we simply collect them from the literature.
- **We do not** judge whether or not the suggested rules from the literature are indeed necessary or sufficient to produce a good colormap, but only catalog them objectively.
- **We do not** claim that all criteria for the design of a good colormap can be put into a mathematical framework; we only assign formulas when it is possible.
- **We do not** consider information about the data, user, display, or visualization task, using a general framework assuming only the colormap itself.
- **We do not** restrict our framework to a specific colorspace, hypothesizing instead the existence of a metric space which describes human color perception.
- **We do not** claim that the suggested measures are unique or the best ones available.
- **We do not yet** provide a tool for colormap improvement. Our hope is that the work will provide a foundation that will eventually lead to automated approaches to colormap improvement.

This common framework has two important applications. First, it will enable scientists of different fields, users, and artists to use a common language when discussing properties of a colormap. Secondly, it will enable the development of algorithms for the automatic assessment, improvement, as well as generation of colormaps. We have implemented the mathematical measures in a publicly accessible online tool at <http://colormeasures.org>. Users can determine the measures of their own colormaps by simply loading the corresponding .json files.

2 RELATED WORK

We first consider the breadth of colormap design rules and suggestions available in the literature with the goal of gathering relevant terms and meanings. We build on excellent survey works of the literature [68, 69, 90] and summarize chronologically. Please keep in mind that many of the rules have been intuitively applied for hundreds of years [1, 62, 70, 81]. For the reader wishing to review general color foundations first, please see Section 3 before returning to this literature review.

In the context of computer graphics, Sloan and Brown [71] mention as early as 1979 that a colormap should consist of colors that are *maximally distinguishable* and have an *order* that can be remembered easily. They suggest treating the colormap as a path through color space and mention the dependence on task, human perception, and display.

Design Rule	Source
Order (intuitive, natural, easy to remember)	[8, 15, 29–31, 46, 50, 52, 55, 71, 79, 81, 83, 86]
Discriminative power (separation, sensitivity, just noticeable differences, distinct color levels, color space utilization/ exploitation, perceptual range / resolution, discriminability)	[15, 30, 41, 42, 46, 50–52, 55, 57, 59, 71, 74, 79, 89]
Uniformity (equidistant differences, separation, associability, separability, linearity, equal values shall be mapped to equal colors)	[8, 15, 19, 29, 30, 40–42, 46, 50–52, 57, 60, 74, 76, 78, 79, 82, 86, 89]
Smoothness (continuity, no boundaries, no Mach bands, low curvature no sharp bends)	[6, 8, 29, 30, 46, 55, 57, 60, 74, 78, 86]
Equal visual importance	[6, 31, 42, 55]
Robust to vision deficiencies	[31, 46, 55, 88]
Robustness to contrast effects	[55, 83]
Robustness to shading on 3D surfaces	[40, 46, 55]
Background sensitivity	[6]
Device independence (do not leave the gamut)	[6, 40, 46, 76, 86, 88]
Aesthetically pleasing	[39, 40, 46, 88]
Intuitive / natural color choices	[64, 77]
Use different colormaps for different variables	[77]
Separation of values into low, medium, and high	[46, 55]
Avoid rainbow	[8, 46, 57, 60, 81]
Highlighting of prominent values	[76]

Table 1: Summary of the suggested perceptual colormap design rules from the literature. We provide mathematical formulations that hold in a non-Euclidean metric space for the bold ones.

Wainer and Francolini [81] stress *order* within a retinal variable, and point out the difficulty of ordering color solely by hue.

Meyer and Greenberg [40] advocate already in 1980 *uniformity* in a *perceptual colorspace*. They further suggest that the path of the colormap through a perceptual color space may not leave the *gamut* of a device, but should be close to the boundary to have brilliant colors. They also suggest to only vary in hue for the coloring of 3D surfaces to *not interfere with the shading*.

Trumbo [79] states that a good univariate colormap must satisfy two fundamental principles, similar to those in [71]. The first is again *order* in one or more retinal variables. However, he explicitly includes hue as one of the retinal variables, in addition to saturation and brightness. His second principle is *separation*, meaning that two different values should be represented by perceivably different colors.

Pizer [51] stresses the importance of *equal changes* in the data value to be equally perceivable in its color representation in the context of display devices. He defines a linear device as one for which the curve of just noticeable differences (JNDs) is constant. In [50, 52], Pizer et al. require a colormap to satisfy *naturalness of order*, *sensitivity*, and *associability*. Sensitivity in this sense corresponds to the length of the path of the colormap through colorspace measured in the number of JNDs, and associability to the fact that close values are mapped to similar colors and more distant values to more different colors. They identify two criteria that guarantee associability: *continuity* and *monotonicity* in brightness. In their opinion, natural order can be achieved by a monotonic increase in brightness and each of the RGB components, such that the order of their intensities does not change throughout the colormap. This is satisfied by the heated body colormap but not by the rainbow. They differentiate perceptual tasks: qualitative, referring to interpretation of the overall form of the data, and quantitative, referring to the ability to read or compare exact values.

Tajima [74] states that the best colormaps have paths with *large color differences* and are *perceptually uniform*. He suggests using colormaps with regular color differences in a perceptually uniform colorspace to produce perceptual uniformity.

Robertson and O’Callaghan [57] refer to Trumbo’s rules, making use of perceptual color spaces to implement order and separation. To our knowledge, they were the first to describe that *straight lines* in a perceptual color space produce colormaps of perceptual uniformity. They also note that bent curves can have a greater *color space utilization* advocating paths that are *smooth* and with *low curvature*.

Mackinlay [37] develops an automated graphical design tool with an algebraic underpinning.

Mathematical Rule	Source
Monotonicity in an attribute (luminance, RGB, saturation and hue, CIELCH)	[5, 8, 19, 25, 30, 31, 49, 50, 52, 55, 58, 60, 61, 72, 83, 88]
Invertibility	[6, 30, 78]
Continuity	[6, 8, 29, 30, 46, 55, 57, 60, 74, 78, 86]
Linearity	[30, 46, 57, 72, 76, 78]
Constant speed	[8, 51]
Long path	[50, 52]
Low curvature	[57]
Redundancy (invertibility / monotonicity in more than one attribute)	[19, 55, 69, 72, 83]
Non-monotonicity in a color-opponent channel	[55, 83]
Fix order of magnitudes of RGB	[50, 52]

Table 2: Summary of suggested mathematical rules from the literature.

Ware [83] points out that a *monotonic change in luminance* is important to see the overall form of the data (qualitative task). On the other hand, he stresses the significance of *non-monotonicity* in at least one color-opponent channel. His experiments show that a colormap that consists of only one completely monotonic path in a single perceptual channel is prone to error in reading the exact values of the underlying data (quantitative task) due to the simultaneity effect.

In an early algorithmic approach, Pham [49] produces *low curvature* colormaps by fitting splines through given points in color space.

Levkowitz and Herman [30] require *order*, *uniformity*, and *no perceivable boundaries*. They define uniformity such that equally spaced data values are represented by colors that are perceived as equally different. This unifies Pizer's definitions of associability and separation and implies linearity. They suggest an algorithm that produces colormaps that have maximal color differences, are monotonic in RGB, hue, saturation, and brightness. In [29], Levkowitz suggests an algorithm to create colormaps with equal speed between adjacent points.

Drawing on expertise and experience in cartography, Brewer [10, 11] advocates hue, lightness, and saturation as the perceptual dimensions and suggests avoiding confounding attributes. Similar to the concept of associability, she demands that progression along a perceptual attribute should relate to progression in data values.

Bergman and Rogowitz, et al. [5, 58–61] distinguish different tasks, data types, and spatial frequency, recommending colormap properties for each combination. They mention that *equal visual importance*, *perceptually even spacing*, *smoothness*, and *monotonically increasing luminance*, saturation, or hue are important for the isomorphic task, which has the goal of faithfully reflecting the structure of the underlying data. They reject the rainbow colormap for failing at this task.

Rheingans [55] stresses the importance of considering the characteristics of the data, the goals of the visualization, and the audience. She summarizes many of the previously suggested rules and adds *robustness w.r.t. color vision deficiency (CVD)*, *classification into low, medium, and high values*, and *little interference with 3D shading*. She also introduces the taxonomy of *redundant* colormaps, in which the information is encoded on more than one attribute.

Light and Bartlein [31] discuss the rainbow colormap and its lack of robustness with respect to CVD. Borland and Taylor [8] also focus on the flaws of the rainbow. A graphical example of a thought experiment drawn from Ware [84] considers the problem of ordering four colors drawn from the spectrum, thus demonstrating its lack of perceptual ordering. The tendency of the rainbow to alternately obscure features in the data and create artifacts is also shown.

Schulze-Wollgast et al. [67] focus on the comparison task. They extract statistical information from the data, e.g., minimum, maximum, average, median, mode, skewness, and quartiles and adjust the colormap to gain a better color discrimination.

Zhang and Montag [89] construct colormaps in CIELAB, evaluating their performance via user studies. They stress the importance of *perceptual uniformity* and *color space exploitation*: the number of distinct color levels through which the path passes.

Tominski et al. [78] are the first to explicitly relate the *invertibility* of the colormap to its effectiveness. They also demand *associability*

and *perceptual linearity* and stress that the characteristics of the data, tasks, goals, user, and output device need to be taken into account.

Wijffelaars et al. [86] state that required properties for colormaps are *perceived order*, *equal perceived distances*, and *equal importance*. They say that the latter is violated if the path of the colormap through space has sharp bends.

Moreland [46] demands a colormap to be *aesthetically pleasing*, have *maximal resolution*, minimal interference with *shading on 3D surfaces*, *robust to CVD*, *order*, *perceptual linearity*, and *not leave the gamut*. He provides a mathematical definition of perceptual uniformity in a local as well as a global sense. He presents an algorithm for the construction of diverging color maps that have a long path through CIELAB and have *no Mach bands* stemming from non-smooth bends in the colormap path. His cool/warm diverging colormap has replaced the rainbow as the default colormap in ParaView [2].

Gresh [19] measures the perception function, which shows how big the difference in color needs to be for a given point in the colormap such that a certain user on a certain monitor can perceive a difference. She goes on to develop an algorithm to transform it into a constant function in order to achieve *equal perceptual steps* in the colorscale.

Thompson et al. [76] suggest applying special colors outside the usual gradient of the colormap to dominantly occurring values.

Mittelstädt et al. [41, 42] require *perceptual linearity* but also a *high discriminative power* and state that saturated colors are important to achieve the latter. They publish quality measures to evaluate how well colormaps preserve data relations based on the stress [75] and how well they exploit the color space using the volume of the convex hull of all used colors in CIELAB or the number of JND's, and the visual importance using the arc of intensity and saturation.

Bernard et al. [6] suggest definitions of colormap properties and build relations to mathematical criteria for their assessment and map them to different tasks in the context of bivariate colormaps. They name *color exploitation* (number of JNDs), *separability*, *background sensitivity* (JND distance to black or white), *device independence*, and *ease of implementation*, and distinguish *perceptual linearity* in a local from a global sense. For the measurement of how perceptually linear a colormap is, they use the variance of the different slopes.

Fang et al. [17] provide an algorithm for optimizing distances between multiple discrete colors.

Experiments by Padilla et al. [48] show that binning a colormap usually leads to longer response times, but more accurate results for a variety of tasks on 2D scalar fields.

Thyng et al. [77] provide a set of colormaps for ocean data. They agree that *uniformity* is important and suggest two new rules. *Consistency* implies that within the same context, two variables should not be visualized by the same colormap just as two variables should not be represented by the same Greek symbol. *Intuition* means that cultural implications and the nature of matter can improve understanding.

Samsel et al. [64] also make use of *intuitive* colors for the visualization of environmental data. They provide sets of blue colormaps for water, browns for land and greens for vegetation. Using the natural colors of different matter aims at exploiting automated cognitive processes that require less conscious concentration [4].

Please note that not all rules are advocated for as hard and fast by the authors. They are often considered beneficial qualities between which a trade-off needs to be found. At times, design rules are explicitly stated w.r.t. a specific task [3]. A large body of research is dedicated to properties of colormapping that can only be evaluated if the data is known, because certain perceptual effects depend on the frequency, the size of a color patch [16, 73], or the composition of its surroundings [43, 45]. Other colormap qualities can not be judged without knowing the task or goal of a visualization [3, 55, 71, 78, 83]. Also the audience, their experience, cultural background, language, names given to colors, personal preferences, aesthetics, and intuitive associations can influence interpretation of a visualization [21, 32, 65, 66]. We mention these aspects to give a more complete overview on the literature, but in this paper, we concentrate on deriving a general framework that does not depend on the data, audience, task, display, goal, or colorspace. This is possible, because we often do not need to take the context into account in order to

define a rule. For example, Ware [83] states that monotonicity supports the qualitative task, while impeding the quantitative one. However, we do not need to know any of that to assign a formula to monotonicity.

Using the nomenclature from this paper, we have collected and summarize the suggested rules from the literature. They fall into three main categories: perceptual, mathematical, and operational rules. The perceptual rules are fundamental. Table 1 is our attempt to gather the most important ones from the literature. In contrast to the perceptual rules, the mathematical ones are merely auxiliary tools to achieve the former. They already come with unambiguous definitions. The most common ones from the literature are collected in Table 2. Finally, the operational rules, such as the use of a perceptually uniform colorspace [29, 30, 40, 49, 57, 72, 88, 89], device independence and adherence to display gamut [6, 40, 46, 76, 86, 88], or ease of implementation [6] are practical rules. They form the groundwork on which the mathematical rules can be applied. Last but not least, even though not always stated explicitly, the underlying motivation of each of these suggested rules is the generation of a good visualization. That means they all follow the highest rule of producing images that best enable the observer to understand the underlying data.

In this paper, we will define a coherent taxonomy for the most common or important perceptual rules. This will enable us to match the bold perceptual rules to a mathematical counterpart, providing them with an unambiguous definition.

3 FOUNDATIONS

We start with a recap of the necessary theoretical foundations.

3.1 Colorspaces

The widely used base for color measurement is the Commission Internationale de L'Eclairage (CIE) XYZ system. It is based on the concept that all humans have a set of three cone receptors with the same color sensitivity functions. Although these receptor sensitivities have been estimated using RGB primaries, the color standard uses the imaginary primary colors X, Y, and Z, which enables all perceivable colors to have positive, device independent coordinates [12]. However, CIEXYZ is designed for color measurement and only relates very indirectly to color appearance. It is perceptually non-uniform.

Uniform color spaces have been developed that attempt perceptual uniformity in the sense that equal metric differences in the spaces should correspond to equally large perceived difference between pairs of colors. The seminal work on defining perceptual color differences was done by Wright and Pitt [87] and MacAdam [36] ("MacAdam ellipses"). Many authors agree that a perceptually uniform color space should be used to assess the quality of colormaps [29, 30, 40, 49, 57, 88, 89]. Examples are the CIE standards, CIELUV, CIELAB, and CIECAM02-UCS. However, these too suffer from deficiencies in that experiments have shown that perceived color differences cannot be captured in a Euclidean space [23, 53]. Several metrics such as ΔE_{1994} or ΔE_{2000} have been proposed on top of CIELAB to achieve greater perceptual uniformity, but these cannot be visualized in a 3D Euclidean space such as the popular chromaticity diagrams; the resulting space does not have an inner product or norm associated with it. Thus there is no straight forward definition of an angle. In addition, even these metrics are not very accurate for the evaluation of large color differences [38]. They may produce discontinuities for colors whose hue angles differ by 180° . In certain areas they do not even satisfy the triangle equation, mathematically disqualifying them as a metric. For example, in ΔE_{2000} , white and grey as well as black and grey differ by $\Delta E_{2000} = 36$ each, but white and black differ by $\Delta E_{2000} = 100$.

Another problem shared by even the best metrics is that they do not account for effects such as simultaneous color contrast and the discounting by the eye and brain of perceived illumination [27], which can result in large distortions in color space. CIECAM is a color appearance model [16] designed to estimate and counterbalance some of the interdependencies of the appearance of a color and its surroundings [35, 47].

In this paper, we are concerned with general purpose color sequences. Therefore, for our theoretical work, we will assume the existence of

a hypothetical metric ΔE , that does perfectly represent the perceived distance between two colors. As practical examples, we evaluate the measures in the most commonly used metrics, ΔE_{1976} , ΔE_{2000} [34], and $\Delta E_{CAM02-UCS}$ [33], henceforth denoted by ΔE_{76} , ΔE_{00} , and ΔE_{02} respectively. By making only this single assumption, our distance metric is an updatable module.

3.2 Colormaps

The basic data in scientific visualization are scalar fields $f: \mathbb{R}^d \rightarrow \mathbb{R}, p \mapsto f(p)$. They associate a real value f , describing a physical quantity, with each point p in space. The most popular visualization of two-dimensional scalar fields is color coding, where each point is assigned a color to represent the corresponding physical value. A continuous colormap forms a curve in a colorspace.

Definition 1 A colormap is a function $x: [a, b] \subset \mathbb{R} \rightarrow C$, which is defined by a colorspace C , an increasing sequence of sampling points $t_0 < \dots < t_m \in [a, b]$, a series of values in the colorspace $x_0, \dots, x_m \in C$, the mapping $x(t_i) = x_i, i = 0, \dots, m$, and a rule for interpolating the intermediate values $t_{i-1} < t < t_i \in [a, b]$.

Unless stated otherwise, we will always interpolate linearly in CIELAB equipped with the Euclidean metric ΔE_{76} in this paper. As implemented in many tools, this makes a colormap a polygonal chain in CIELAB. It is important to note that a straight line from linear interpolation generally does not coincide with the shortest path if we use a non-Euclidean metric such as ΔE_{00} . This can be seen if we compare Figures 7(a) and (b). The greyscale and cool/warm divergent colormaps are uniform with respect to ΔE_{76} , where they were designed, but not with respect to ΔE_{00} .

Colormaps span a range of types including continuous, cyclical, discrete, banded, and categorical [5, 11, 42, 46]. In this work, we restrict the development of the theoretical framework (Section 4) to continuous colormaps. However, many of the results can either be directly applied to other types of colormaps or applied via extrapolation of the methodology. Please note that Definition 1 includes discrete colormaps if we use a constant interpolation scheme instead. Cyclical colormaps can likewise be addressed with either linear or constant interpolation as appropriate while using an angular difference rather than linear difference for the sampling points. The half greyscale and flat greyscale showcase maps are indicative of the potential order issues that one would also expect with banded and cyclical colormaps. Divergent colormaps, often most appropriate for interval data, are a subset of continuous colormaps that can also be assessed with these measures, as shown by the cool/warm colormap included in our showcase set. Some divergents may exhibit non-constant local speed in the vicinity of the divergent point, but application of our framework is certainly appropriate to divergent colormaps. A more rigorous extrapolation of this work to the broader range of colormap types must be left for future work.

In order to express the measures for continuous functions, we would need infinitesimal mathematics. However, to keep the notation easily accessible and readily implemented, we will assume that we have a sufficiently fine, equidistant sampling on the unit interval $t_j = j/n, j = 0, \dots, n$ for each colormap and use corresponding discrete formulations. This resampled version of a colormap can be derived using the above interpolation. The 2D images and measures in this paper were generated using $n = 20$.

3.3 Available Measures for Colormap Assessment

One of the goals of this paper is to find relationships between the perceptual rules in Table 1 and the mathematical ones in Table 2. But as we discussed previously, this will not always be possible or reasonable. For example, it seems intuitive that the perceptual quality of smoothness is related to the curvature, but a metric space does not necessarily provide the concept of curvature. Therefore, we will restrict ourselves to quantities that can be measured.

In our framework, there is a distinction between the definition of a rule and the measure for its evaluation. While the definition is

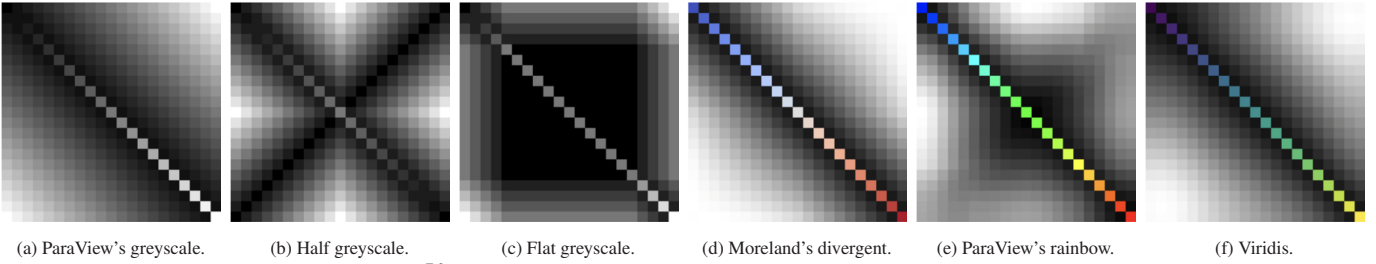


Fig. 2: The matrix of the **global distances** $D_{i,k}^{76}$. The colormaps are displayed on the diagonal $i = j$. The luminance of each patch $i \neq j$ represents the mutual distance between the two colors at the points $x(t_i)$ and $x(t_k)$, which can be seen going horizontally and vertically towards the diagonal.

meant to be ubiquitous, there may be several measures that evaluate a given property. For example, the stress [24, 41, 63] can be a measure of the uniformity, as can the standard deviation of the speed, or the acceleration. Also, measures may allow for an informal comparison in the sense that colormap A may satisfy a rule more or less than colormap B even if the rule can only assume the values true or false.

We will see in the next section that the (bold-faced) concepts in Table 2 can all be evaluated from derived quantities of the global distances. Please note that we do not claim that these measures are the best ones possible. A detailed experimental comparison is beyond the scope of the current paper and must be left for future work.

To guarantee that our framework is as general as possible, it must be independent of the data, the task, and the colorspace. Therefore we are limited to a single measure: the perceived distance between two colors in the colormap. This is all that the assumed metric, ΔE , provides. In the following section, we will use this measure to define the different design rules.

For a colormap, the mutual **distances** between all pairs of points $t_i, t_j, i, j \in \{0, \dots, n\}$ give a 2D scalar function $D_{i,j} : [0, 1]^2 \rightarrow \mathbb{R}$

$$D_{i,j} := \Delta E_{ij} = \Delta E(x(t_i), x(t_j)) \quad (1)$$

as illustrated in Figure 2. The matrix visualization, similar to Demiralp et al. [14], uses luminance to indicate the mutual distances between any two colors. Higher luminance indicates a greater distance. From this, we can derive various statistical quantities such as minimum, maximum, mean, and standard deviation.

Additionally, the distance allows for a definition of the metric **speed** $\forall i \neq j \in \{0, \dots, n\} : V_{i,j} : [0, 1]^2 \rightarrow \mathbb{R}$,

$$V_{i,j} := \frac{D_{i,j}}{|t_j - t_i|}, \quad (2)$$

shown in Figure 1 for the showcase colormaps. Note that in a normed vector space, we could use the velocity, i.e. the vector valued derivative of the distance, but in an arbitrary metric space, it does not exist.

3.4 Local and Global Measures

As briefly touched on in Section 1, we found in the review of the literature that when noticeably different concepts were referred to with the same terms, they could often be decoupled by distinguishing a local interpretation from the global interpretation. Robertson [56] notes: “The significance, and exploitation, of the difference between local interpretation and global interpretation (or concentration) is particularly clear, for example in the graphic works of M. C. Escher, the music of J. S. Bach, and in many other artistic fields [22].” We will likewise use this distinction of local and global quantities for our framework.

Global quantities describe the relationship of a color in a colormap to any other of its colors. The two-dimensional function of **global distances** is given in (1) and the **global speeds** in (2). In contrast, the local concepts describe the relationship of a color and its near neighbors.

In this sense, we can derive the one-dimensional **local distance** function $d_j : [0, 1] \rightarrow \mathbb{R}$ between two neighboring points on the colormap $\forall j = 1, \dots, n$:

$$d_j := D_{j,j-1} \quad (3)$$

and analogously the **local speed** $v_j : [0, 1] \rightarrow \mathbb{R}$, illustrated in Figure 7:

$$v_j := V_{j,j-1}. \quad (4)$$

4 THEORETICAL FRAMEWORK

In this section, we analyze the most common design rules from the summary in Table 1: discriminative power, uniformity, and order. For each concept, we will first browse the related work for nuances in the way they are used by the different authors, then specify the taxonomy. Finally, we will provide an unambiguous mathematical definition using only the measures of distance and speed introduced in the previous section and their derived quantities such as minimum, average, or standard deviation. The restriction to these measures for the evaluation of the design rules is crucial for the generality.

We will illustrate the quality measures using a set of six showcase colormaps shown in Figure 1. The greyscale and rainbow are from ParaView [2], cool/warm divergent from Moreland [46], and Viridis [80] from Matplotlib. The adaptations of the greyscale we constructed ourselves for illustrative purposes. We will also briefly discuss smoothness, attribute-related measures, and robustness, but without mathematical formulations.

4.1 Discriminative Power

We decided to use the terminology of the **discriminative power** of a colormap to unify the terms *maximally distinguishable* [71], *separation* [55, 79], *sensitivity*, or *dynamic range* [30, 50–52], *color differences*, *number of distinct colors*, *number of just noticeable differences* [30, 74], *color space utilization* [57], *color space exploitation* [6, 15, 41], *number of distinct colors/color levels* [6, 15, 41, 42, 50–52, 89], *perceptual range*, *perceptual resolution* [46], *discriminability* [15], because the way that these different terms have been used suggests a common idea.

Pizer et al. [50–52], describe how it can be evaluated experimentally using the notion of *just noticeable differences* (JNDs) and Rogowitz et al. [59] demonstrate a psychophysical experimental methodology. From the way, they and others [57, 74] discuss discriminative power, it corresponds to the arclength $\sum_{j=1}^n d_j$ of the colormap’s curve through colorspace. We will call this interpretation the **local discriminative power** of a colormap as it only takes the relationship between neighboring points into account. On the unit interval, it coincides with the **average local speed** $\bar{v} \in \mathbb{R}$:

$$\bar{v} := \frac{\sum_{j=1}^n v_j}{n-1} = \frac{\sum_{j=1}^n v_j h}{\sum_{j=1}^n h} = \frac{\sum_{j=1}^n v_j (t_j - t_{j-1})}{\sum_{j=1}^n t_j - t_{j-1}} = \frac{\sum_{j=1}^n d_j}{(t_n - t_0)} = \sum_{j=1}^n d_j. \quad (5)$$

Other authors describe the discriminative power in a broader sense [15, 41, 42, 55, 71, 79, 89]. They refer to a more global meaning in which not only the distances between neighboring colors play a role but across the entire colormap. We will call this interpretation the **global discriminative power** of a colormap. The straightforward generalization of arclength to a two-dimensional function would be the surface area, but its computation would require the existence of a cross product, which does not necessarily exist in a metric space. Summing up all pairwise distances $\sum_{i \neq j=0}^n D_{i,j}$ is also not very helpful because it grows with the number of sample points n . A measure that can be

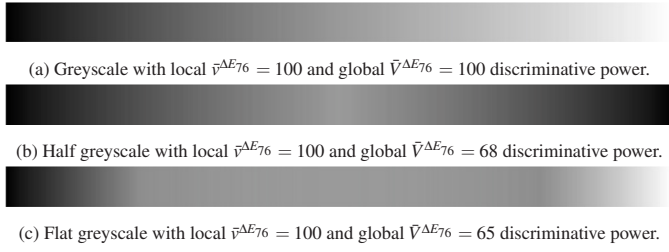


Fig. 3: All colormaps have the same local but different global discriminative powers.

directly generalized to measure the global discriminative power is the **average global speed** $\bar{V} \in \mathbb{R}$:

$$\bar{V} := \frac{\sum_{i \neq j=0}^n V_{i,j}}{n(n-1)}. \quad (6)$$

An illustration of the difference of the two measures can be found in Figure 3. The average global speed of our showcase colormaps can be seen in Figure 4.

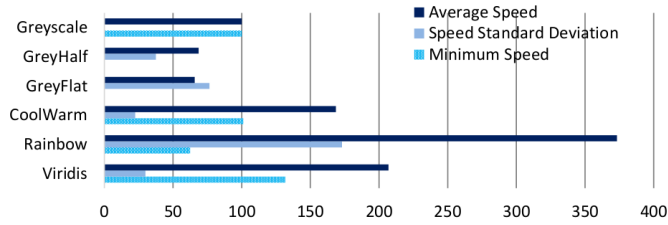


Fig. 4: The **global speed** characteristics of the showcase colormaps. A colormap's average speed correlates to its discriminative power. It is highest for the rainbow and lowest for the greyscales. The minimum speed measures its legend-based order. It is violated for GreyHalf and GreyFlat. The standard deviation of the speed measures its uniformity. The Greyscale is uniform in ΔE_{76} .

4.2 Uniformity

We chose the term **uniformity** [30, 40, 57] to encompass the terminology of *separation* [55, 79], *associability* [50–52], *equal spacing* [57] *perceptual linearity* [6, 50–52, 78, 88], *even spacing* [60], *representative distance* [30], *perceptual uniformity* [46, 57, 74, 82, 89], *equal perceived distances* [86], *equal perceptual steps* [19], *preservation of data relations* [10, 11, 41, 42], *separability*, and *preservation of data distances* [6], all of which appear to follow a common concept.

The definition of uniformity in the literature is not very controversial. To our knowledge, the first explicit formula was given by Levkowitz and Herman in 1992 [30]. While some authors seem to refer to uniformity in a global sense [55, 79], others only consider neighboring sample points [74]. Many are aware of this difference and make a distinction between local and global uniformity [6, 41, 42, 57]. The difference is illustrated in Figures 5 and 6.

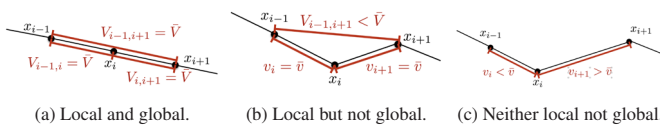


Fig. 5: Illustration of the uniformity, which corresponds to the standard deviation of the speed. For local, the colormap must suffice $\sigma v = 0$, for global, $\sigma V = 0$.

One approach to experimentally test uniformity would be to present the observer with two pairs of colors. A colormap is uniform if the

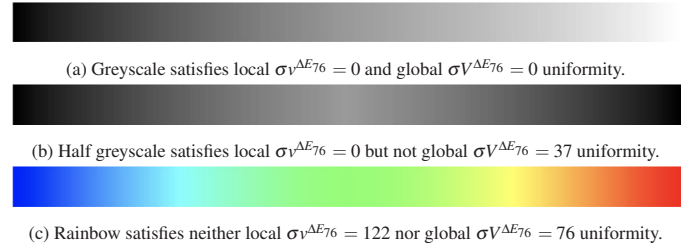


Fig. 6: Uniformity of different colormaps. For local, they must suffice $\sigma v = 0$ and for global $\sigma V = 0$.

observer perceives the pair as more similar for those pairs that correspond to sample points that are less distant. An example can be found in [86]. For the local interpretation, the pairs need to be formed from three consecutive colors.

A colormap suffices **local uniformity** if the distances between adjacent colors correspond to the distances of the values they represent:

$$\forall j \in \{1, \dots, n\} : \frac{d_j}{t_j - t_{j-1}} = \frac{d_{j+1}}{t_{j+1} - t_j} \quad (7)$$

It is obvious from the definition that the local uniformity is equivalent to constant local speed. The local uniformity of our showcase colormaps is shown in Figure 7. The less the local speed deviates from the average, the more uniform is the colormap. For an overall measure of local uniformity, we can use the **standard deviation of the local speed** $\sigma v \in \mathbb{R}$:

$$\sigma v := \sqrt{\sum (v_j - \bar{v})^2} \quad (8)$$

Low standard deviation corresponds to high local uniformity.

A colormap satisfies **global uniformity**, if

$$\forall i, j, k, l \in \{1, \dots, n\} : \frac{D_{i,j}}{|t_j - t_i|} = \frac{D_{k,l}}{|t_l - t_k|}, \quad (9)$$

which is equivalent to the global speed being constant. Global uniformity is a very strong constraint. In a Euclidean setting, it would be equivalent to the mathematical definition of **linearity**, which means, using the ΔE_{76} metric, that only straight lines in CIELAB fulfill it. This is probably the reason why Levkowitz and Herman [30] and also Moreland [46] define uniformity in a global sense, but enforce only a local version when constructing their colormaps. In a non-Euclidean setting, linearity is not defined in a straightforward way.

Analogous to the local case, an overall measure of global uniformity is the **standard deviation of the global speed** $\sigma V \in \mathbb{R}$:

$$\sigma V = \sqrt{\frac{\sum_{i \neq j=0}^n (V_{i,j} - \bar{V})^2}{n(n-1)}}, \quad (10)$$

which can be found in Figure 4 applied to the six showcase colormaps. As expected, it is zero for the linear greyscale and maximal for the rainbow.

4.3 Order

The term order is sometimes used to describe different properties of a colormap. Sloan and Brown [71] use order in the sense that it can be remembered easily. Thus they would allow a user to consult the legend before ordering the colors. In contrast, other authors [46, 52, 57, 83] demand a natural or intuitive order, such that a user should not need a legend to be able to order the colors. Levkowitz and Herman [30] give it a local meaning. They want a color x_{i-1} to perceptually precede its successor, i.e. $x_{i-1} < x_i, \forall i = 1, \dots, n$. On the other hand, Ware [84] refers to order in a global sense. Not only adjacent colors, but colors picked from anywhere in the colormap should be sortable by a user. While Ware does not distinguish whether the user sorts them from low to high or vice versa, Moreland wants the user to be able to associate

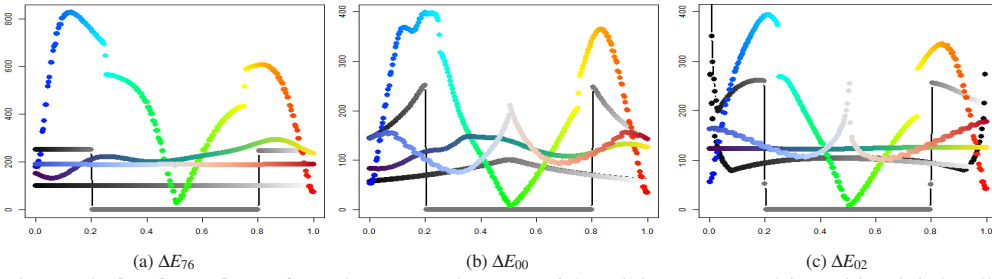


Fig. 7: The **local speeds** v_j of our showcase colormaps. If the minimum $v_{min} > 0$ is positive, it is locally invertible, i.e. we have **local legend-based order**. The flat greyscale has a local speed of zero between 0.2 and 0.8, because it is **identically grey** there. All other colormaps can be locally ordered. The higher the average local speed \bar{v} , the higher is the **local discriminative power**. It is equally low for all greyscales and highest for the rainbow. The deviation from the average is a measure of the **local uniformity**.

low and high values with the colors on the colormap [46]. Taking all of this into account, we define different notions of order as follows.

Experimentally speaking, a colormap is considered to satisfy **order** if for a given a set of colors picked from the colormap, any general observer would sort them in the same way [84]. We add the following specifications to account for the different uses of the term in the literature and to make them distinguishable.

Intuitive vs easy to remember vs legend-based: Either the sorting is performed without, or after, or during access to the legend.

Local vs global: The sample points are chosen either consecutively or arbitrarily.

Directed vs undirected: Either the user can sort two colors in the same order, or is able to consistently pick the middle one out of three colors.

Please note that legend-based order implies directed order. Since a metric by definition is symmetric, it is not possible to describe the concept of directed, intuitive order per se. The concept of an easily remembered order is strongly dependent on the cognitive abilities of a human and cannot be expressed in our metric space. We treat the remaining cases in the following subsections.

4.3.1 Legend-Based Order

Local legend-based order is related to the concept of **local invertibility**. It is fulfilled as long as two neighboring colors do not coincide:

$$\forall j \in \{1, \dots, n\} : x_{j-1} \neq x_j. \quad (11)$$

We could use the minimum local distance, $\min_{j=1}^n d_j$ and check if it is non-zero. But for human observers, it is not sufficient for the points to be different, they need them to be noticeably different. The intuitive way to evaluate how different neighboring colors are is to use the local distance d_j , but its behavior is not easy to interpret. It depends on the resolution with which we compute it and will decrease with growing n . Therefore, we will use the **minimum local speed** $v_{min} \in \mathbb{R}$:

$$v_{min} := \min_{j \in \{1, \dots, n\}} v_j. \quad (12)$$

It is zero *iff* the colormap is not invertible. The higher it is, the more easily the colormap can be inverted locally. As a side note: since the

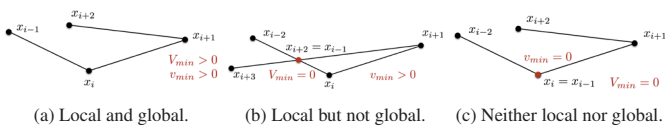


Fig. 9: Illustration of the legend-based order, which corresponds to the invertibility. For local, the colormap must suffice $v_{min} > 0$; for global $V_{min} > 0$.

local speed is the finite approximation to the first derivative, demanding it to be non-zero is the discrete equivalent of the inverse function theorem. Local non-intuitive undirected order corresponds to invertibility in a neighborhood.

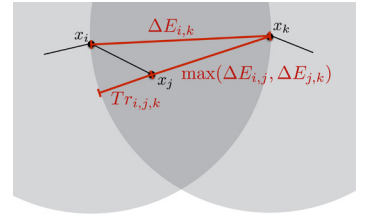


Fig. 8: Illustration of the triangle side difference $Tr_{i,j,k}$. Here, the colors can be correctly ordered, because the longest side connects the two outer points, i.e. the middle point lies within the dark grey area.

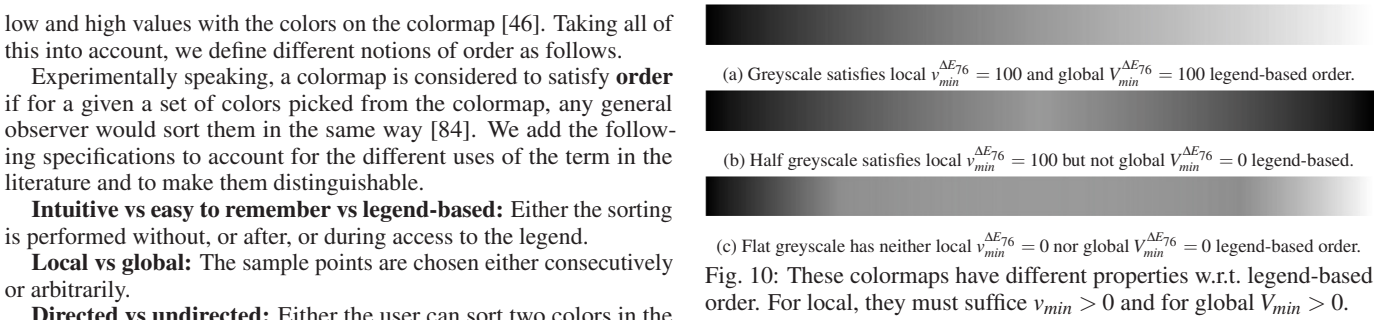


Fig. 10: These colormaps have different properties w.r.t. legend-based order. For local, they must suffice $v_{min} > 0$ and for global $V_{min} > 0$.

Global legend-based order is the analog when the distances do not correspond to neighboring points, but can be taken from anywhere in the colormap. This means it corresponds to the **invertibility** of a function [30, 78]. It is guaranteed if the function is **injective**:

$$\forall j \neq i \in \{0, \dots, n\} : x_i \neq x_j, \quad (13)$$

which means that the colors are mutually distinct.

Again, analogous to the local case, we can measure it using the **minimum global speed** $V_{min} \in \mathbb{R}$:

$$V_{min} = \min_{i \neq j \in \{1, \dots, n\}} V_{i,j}. \quad (14)$$

Examples of the different legend-based orders can be found in Figures 9 and 10.

The overall global legend-based order of our showcase colormaps can be seen in Figure 4. The minimum global speed is zero for the half greyscale and the flat greyscale. Thus these two can not be ordered globally even given a legend. All others can be ordered. Figure 7 shows the local legend-based order. The greyscale with the flat area cannot be ordered. The flat green part in the rainbow exhibits a significant drop in local speed, reflecting the difficulty of ordering in that region.

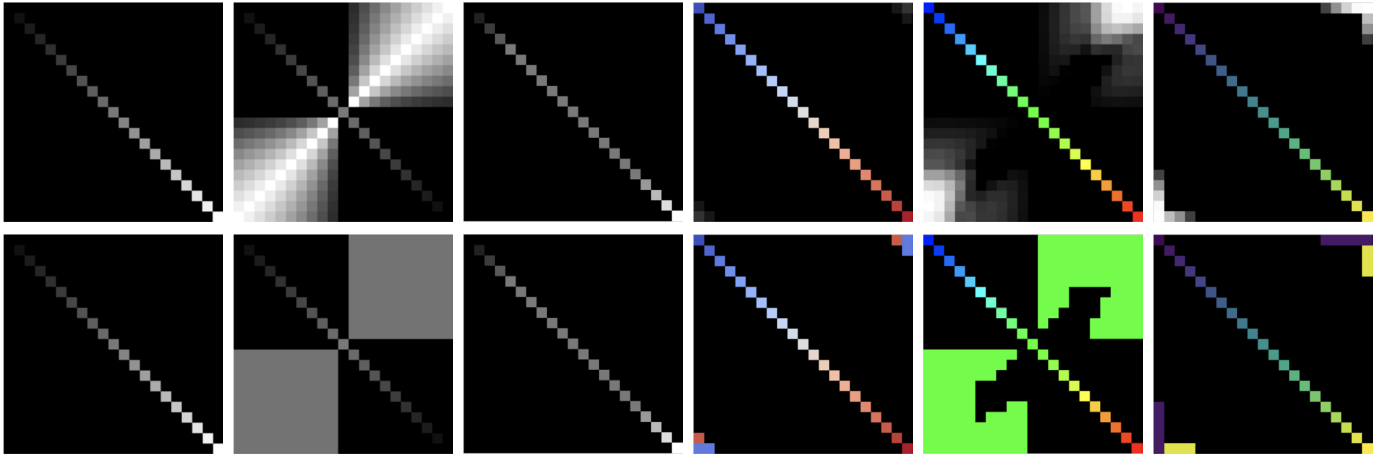
4.3.2 Intuitive Order

Our metric can determine how different an observer perceives pairs of colors. As a result, we expect the observer to correctly sort three colors in an intuitive way if the distance of the two outer colors is larger than their respective distances to the one in the middle. The corresponding mathematical formulation for **local intuitive order** is therefore:

$$\forall j \in \{1, \dots, n-1\} : D_{j-1,j} < D_{j-1,j+1} > D_{j,j+1}. \quad (15)$$

Since the local triangle side ratio $\max(D_{j-1,j}, D_{j,j+1})/D_{j-1,j+1}$ is numerically unstable for small $\Delta E_{j-1,j+1}$ we use the **local triangle side difference** $tr_j : [0, 1] \rightarrow \mathbb{R}$:

$$tr_j := \frac{D_{j-1,j+1} - \max(D_{j-1,j}, D_{j,j+1})}{t_{j+1} - t_{j-1}}. \quad (16)$$



(a) Greyscale $Tr_{\min}^{\Delta E_{76}} = 5$. (b) Half g.s. $Tr_{\min}^{\Delta E_{76}} = -100$. (c) Flat g.s. $Tr_{\min}^{\Delta E_{76}} = -5$. (d) Divergent $Tr_{\min}^{\Delta E_{76}} = -0.2$. (e) Rainbow $Tr_{\min}^{\Delta E_{76}} = -85$. (f) Viridis $Tr_{\min}^{\Delta E_{76}} = -1.4$.

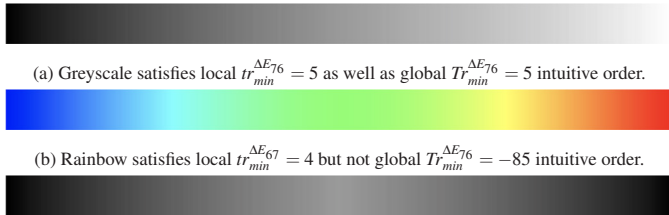
Fig. 11: The **global intuitive order** can be evaluated using the **triangle side difference**. For each (i, k) , the top row shows how much $Tr_{\min}^{\Delta E_{76}} = \min_{i < j < k} Tr_{i,j,k}^{\Delta E_{76}}$ over all j between them goes below zero. The bottom row shows the color that of this j where the minimum is assumed.

It is positive if the colors can be sorted and becomes more negative the further the middle point is away from the outer ones. Scaling with $t_{j+1} - t_{j-1}$ prevents it from decreasing with growing n . An visualization of this concept can be found in Figure 8.

A positive **minimal local triangle side difference** $tr_{\min} \in \mathbb{R}$:

$$tr_{\min} := \min_{j \in \{1, \dots, n-1\}} tr_j \quad (17)$$

indicates that a colormap suffices local order everywhere. The local intuitive order can be read off the subdiagonal of the global plot in Figure 11. It is positive for all of our showcase colormaps except for the point of inflection in the half greyscale.



(a) Greyscale satisfies local $tr_{\min}^{\Delta E_{76}} = 5$ as well as global $Tr_{\min}^{\Delta E_{76}} = 5$ intuitive order.

(b) Rainbow satisfies local $tr_{\min}^{\Delta E_{67}} = 4$ but not global $Tr_{\min}^{\Delta E_{76}} = -85$ intuitive order.

(c) Flat greyscale has neither local $tr_{\min}^{\Delta E_{76}} = -5$ nor global $Tr_{\min}^{\Delta E_{76}} = -37$ intuitive order.

Fig. 12: Intuitive order of different colormaps. For local, they must suffice $tr_{\min} > 0$ and for global $Tr_{\min} > 0$.

The global analogue, **global intuitive order** is given by:

$$\forall i < j < k \in \{1, \dots, n\} : D_{i,j} < D_{i,k} > D_{j,k}. \quad (18)$$

It can be measured by the **global triangle side difference** $Tr_{i,j,k} : [0, 1]^2 \rightarrow \mathbb{R}$:

$$Tr_{i,j,k} := \frac{D_{i,k} - \max(D_{i,j}, D_{j,k})}{t_k - t_i}. \quad (19)$$

If the **minimal global triangle side difference** $Tr_{\min} \in \mathbb{R}$:

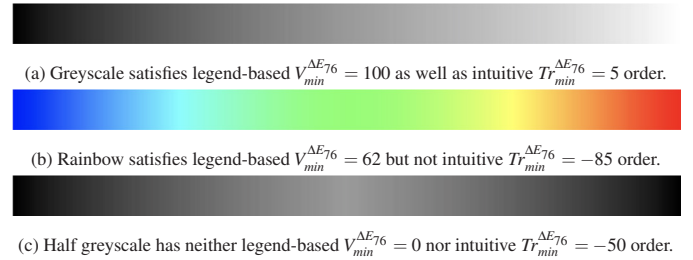
$$Tr_{\min} := \min_{i < j < k \in \{1, \dots, n-1\}} Tr(i, j, k) \quad (20)$$

is positive, the colormap can be intuitively ordered everywhere.

Note that the actual positive value of the triangle side difference itself is not easy to interpret because it assumes values close to zero if one of the points is far away from the other two. That means it does not correlate to how close to perceptually linear a colormap is, but depends on the resolution.

The global triangle side difference is a 3D function $[0, 1]^3 \rightarrow \mathbb{R}$. To visualize it, we check, for every pair i, k , if there is a sample point $j : i < j < k$ with $Tr(i, j, k) < 0$. If so, we plot the j that minimizes $Tr(i, j, k)$ at the coordinates i, k . Thus the user can immediately see which area of the colormaps cannot be globally ordered in an intuitive way and why. From Figure 11, the greyscale can be intuitively ordered. The point of inflection breaks the intuitive order of the half greyscale. In the rainbow, green is furthest away from red and blue.

Examples of colormaps satisfying intuitive order in a global or local sense can be found in Figure 12. The difference between legend-based and intuitive order in a global sense is illustrated in Figure 13.



(a) Greyscale satisfies legend-based $V_{\min}^{\Delta E_{76}} = 100$ as well as intuitive $Tr_{\min}^{\Delta E_{76}} = 5$ order.

(b) Rainbow satisfies legend-based $V_{\min}^{\Delta E_{76}} = 62$ but not intuitive $Tr_{\min}^{\Delta E_{76}} = -85$ order.

(c) Half greyscale has neither legend-based $V_{\min}^{\Delta E_{76}} = 0$ nor intuitive $Tr_{\min}^{\Delta E_{76}} = -50$ order.

Fig. 13: These colormaps have different properties w.r.t global order. For legend-based, they must suffice $V_{\min} > 0$ and for intuitive $Tr_{\min} > 0$.

4.4 Smoothness

Terms such as *no boundaries* [29, 30, 55, 89] *no sharp transitions* [8], *continuity* [8, 46, 50, 52], *no irregular perception* [46], *no artificial Mach bands* [46], *no artifacts* [8, 46, 60, 89], *no sharp bends* [86], or *low curvature* [57, 90] can be grouped into the concept of **smoothness** [46, 57, 60, 79, 88, 90]. This colormap property is surprisingly complex as the perceptual smoothness of a colormap can be influenced by many factors. Areas of a colormap can produce the impression of sharp bends due to varying speed, actual bends in the path of the colormap, changes in visual importance, or borders between areas of colors belonging to a common name group.

There are approaches to mathematically estimate the visual importance or how much attention a color attracts [6, 31, 42, 55]. They are based on how saturated and how bright a color is perceived [13, 20] and are given by $\sqrt{L^2 + \sqrt{a^2 + b^2}}$. For the classical definition of curvature, we need the concept of an angle, not available in a metric space without an inner product. We currently do not see how we can express smoothness through a mathematical formula that is valid without losing the invariance with respect to the underlying metric space or the characteristics of the observer. Development of smoothness measures using only a non-Euclidean metric is left for future work.

4.5 Properties w.r.t. Attributes

Attribute-related properties are defined analogously to the above properties. However, any experiments would be performed after projection of the colormap to the attribute under study. One-dimensional attributes include luminance, hue, saturation, chroma, red, green, blue, the color-opponent channels, or any other arbitrarily chosen one-dimensional submanifold of a colorspace. Two-dimensional attributes are a combination of two 1D attributes, e.g., an isoluminant plane in CIELAB. In this way, we can define order, discriminative power, and uniformity with respect to a specific attribute.

Note that in the literature, certain attributes are sometimes referred to as *retinal* or *perceptual variables* [57, 79, 81], *axes* [25, 57, 74, 88, 89], (*perceptual*) *dimensions* [50, 60, 86, 88], (*retinal/color opponent*) *channels* [55, 83], (*color/primary*) *components* [6, 7, 52, 55, 71, 74, 86], *subspaces* [9], among other terms. We chose the term **attribute** because the other terms are either too restrictive or in conflict with fundamental definitions in mathematics, physics, biology, or perceptual science.

In the literature, the most commonly suggested property of an attribute is **monotonicity**. A one-dimensional function $f : \mathbb{R} \rightarrow \mathbb{R}$ is monotonic, if it suffices:

$$\forall s \leq t : f(s) \leq f(t) \text{ or } \forall s \leq t : f(s) \geq f(t). \quad (21)$$

Please note that monotonicity in an attribute is only defined if the attribute is at least a partially ordered set. Luminance is the attribute most often associated with monotonicity [5, 8, 19, 25, 30, 31, 49, 50, 52, 55, 58, 60, 61, 83], but monotonic behavior has also been demanded for RGB components [30, 50, 52], saturation, hue [30, 55], or the dimensions of CIELCH [88]. Also the explicit violation of the rule, namely non-monotonicity in at least one color-opponent channel has been suggested [55, 83].

The question of **redundancy** [19, 55, 69, 83] also falls into the category of attributes. A colormap is perceptually redundant with respect to a property if it still satisfies the property in question after an attribute is removed. Further properties regarding attributes that have been suggested include a fixed order of the magnitudes of the RGB components [50, 52] or uniformity in only the luminance attribute [28]. Monotonicity is often referred to as a property necessary for order. A rigorous experimental analysis of this will be future work.

4.6 Robustness

Generally speaking, a colormap is robust with respect to any of its defined perceptual properties if it performs comparably well when distorted. The most common sources of distortion include *CVD* [31, 42, 46, 55, 88], *perceptual interaction with the background* [6, 55], *dependence on the display device* [6, 88], *contrast effects* [55, 83], or *interference with the shading on 3D surfaces* [46, 55]. We currently do not see how we can express robustness through a mathematical formula without taking user, task, and data into account. However, for the interested reader, methods such as outlined in Mittelstädt, et al. [44, 45] can be used to compensate for background effects and spatial distributions of the data.

5 COLORMEASURES.ORG

In order to facilitate comparing colormaps and informing decisions on choice of colormaps, the measures for colormap assessment and the visualizations from this paper have been implemented in an online tool at <http://colormeasures.org>. There, users can reproduce our results and assess any arbitrary good, bad, or ugly colormap of their choosing. After opening the colormeasures viewer, a user can browse their computer to choose a colormap (in .json format) or download one of our examples. The colormeasures tool will output the mathematical measures for each of the three colorspace discussed, ΔE_{76} , ΔE_{00} , and ΔE_{02} , as well as the visualizations described above. Additionally, the measure values are output into a form where they can be easily copied offline for further study and comparison. We invite the user to test the colormeasures tool, assess their own colormap choices, and use this information to inform their choice of a colormap to best fit the needs of their data. We welcome any feedback or suggestions for improvement.

6 DISCUSSION AND FUTURE WORK

We have developed a mathematical framework to describe and assess colormap properties, based solely on an assumed ΔE metric that mimics human perception. By separating local from global interpretations of the colormap design rules from the literature, we were thereby able to distinguish different uses of the same terms. This clarification enabled us to use mathematical formulas to cast the taxonomy in stone and provide measures that always exist in any metric space to evaluate them. Often, we could make use of already suggested formulas as we found relationships between the perceptual rules in Table 1 and the mathematical ones in Table 2. Table 3 summarizes our main findings, listing the correspondence between the mathematical and the perceptual concepts as well as our suggested measures for their evaluation.

Mathematical descriptions relating the discriminative power and uniformity to the speed have occasionally been described within the literature. We have structured them into a coherent framework. To our knowledge, this is the first time that a mathematical formulation for the concept of order has been suggested.

We hope that establishing this framework and the resultant measures will aid in experimental design and make colormap assessment more accessible. As an example, a recent paper by Ware et al. [85] develops an experimental approach to assess perceptual uniformity, an approach inspired by both previous research in color and by collaboration on this paper. Using a coherent taxonomy significantly aided the collaboration between researchers from different fields because the necessary terms had unambiguous definitions. The differences between the experimentally measured uniformity from [85] and the predicted measures in the three colorspace discussed here are indicative of how far we have to go to reach a colorspace that truly mimics human perception. Yet the conclusions from Ware et al. are echoed by these mathematically derived measures. Experimental approaches such as [85] and [14], can provide roadmaps to move from theoretical to practical applications.

As discussed in Subsection 3.2, we plan on expanding our framework to discrete and cyclic colormaps in future work. We want to find a mathematical description for smoothness. We plan on conducting experiments to assess whether or not monotonicity in certain attributes implies order. Further, we will collect and compare other measures that can be used to assess the design rules and evaluate experimentally which ones correspond best to human intuition. We also plan on comparing algorithms for the automatic adaption of colormaps to satisfy the suggested design rules, such as uniformization or linearization algorithms. Once we have evaluated which algorithms perform best, we plan on integrating them into colormeasures.org to extend it into an interactive colormap design and improvement tool. While currently agnostic on the use of the suggested measures to declare any specific colormap to be good, bad, or ugly, we hope that by developing these mathematical foundations, this work will lead to further research in color, both experimental and theoretical, that will eventually allow us to provide detailed guidance in choosing and designing colormaps for specific tasks and data types.

Perceptual Rule	Mathematical Rule	Evaluation Measure
L/G discr. power	Long path / -	Average L/G speed
L/G uniformity	Const. speed / linearity	Deviation of L/G speed
L/G Legend-based order	L/G Invertibility	Minimal L/G speed
L/G intuitive order	- / -	L/G triangle side diff.

Table 3: Summary of the relations between the perceptual, the mathematical rules printed bold in Tables 1 and 2, and the available measures in all metric color spaces that can be used for their evaluation. L/G stands for local / global.

ACKNOWLEDGMENTS

We would like to thank Prof. Hans Hagen, Dr. Kenneth Moreland, and Dr. Bernice Rogowitz for inspiring discussions. We thank Curt Canada, Jonas Lukasczyk, and Dr. Jon Woodring for their help. This material is based upon work supported by Dr. Lucy Nowell of the U.S. Department of Energy Office of Science, Advanced Scientific Computing Research under Award Numbers DE-AS52-06NA25396 and DE-SC-0012516.

REFERENCES

- [1] G. A. Agoston. *Color theory and its application in art and design*, vol. 19. Springer, 2013.
- [2] J. Ahrens, B. Geveci, C. Law, C. Hansen, and C. Johnson. ParaView: an end-user tool for large-data visualization. *Visualization handbook*, pp. 717–731, 2005.
- [3] N. Andrienko and G. Andrienko. *Exploratory analysis of spatial and temporal data: a systematic approach*. Springer Science & Business Media, 2006.
- [4] M.-T. Bajo. Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4):579, 1988.
- [5] L. D. Bergman, B. E. Rogowitz, and L. A. Treinish. A rule-based tool for assisting colormap selection. In *Proceedings of the 6th conference on Visualization '95*, p. 118. IEEE Computer Society, 1995.
- [6] J. Bernard, M. Steiger, S. Mittelstädt, S. Thum, D. Keim, and J. Kohlhammer. A survey and task-based quality assessment of static 2D colormaps. In *SPIE/IS&T Electronic Imaging*, pp. 93970M–93970M. International Society for Optics and Photonics, 2015.
- [7] D. Borland and A. Huber. Collaboration-specific color-map design. *IEEE Computer Graphics and Applications*, 31(4):7–11, July 2011. doi: 10.1109/MCG.2011.55
- [8] D. Borland and R. M. Taylor II. Rainbow color map (still) considered harmful. *IEEE computer graphics and applications*, 27(2):14–17, 2007.
- [9] S. Bremm, T. von Landesberger, J. Bernard, and T. Schreck. Assisted descriptor selection based on visual comparative data analysis. In *Computer Graphics Forum*, vol. 30, pp. 891–900. Wiley Online Library, 2011.
- [10] C. Brewer. *Designing better Maps: A Guide for GIS users*. Environmental Systems Research, 2004.
- [11] C. A. Brewer. Color use guidelines for mapping. *Visualization in modern cartography*, pp. 123–148, 1994.
- [12] A. Broadbent. Calculation from the original experimental data of the cie 1931 rgb standard observer spectral chromaticity coordinates and color matching functions. *Québec, Canada: Département de génie chimique, Université de Sherbrooke*, 2008.
- [13] N. Camgöz, C. Yener, and D. Güvenç. Effects of hue, saturation, and brightness on preference. *Color Research & Application*, 27(3):199–207, 2002.
- [14] Ç. Demiralp, M. S. Bernstein, and J. Heer. Learning perceptual kernels for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1933–1942, Dec 2014. doi: 10.1109/TVCG.2014.2346978
- [15] M. Eismann, G. Albuquerque, and M. Magnor. Data driven color mapping. In *Proceedings of EuroVA: International Workshop on Visual Analytics, Bergen, Norway*. Citeseer, 2011.
- [16] M. D. Fairchild. *Color appearance models*. John Wiley & Sons, 2013.
- [17] H. Fang, S. Walton, E. Delahaye, J. Harris, D. A. Storchak, and M. Chen. Categorical colormap optimization with visualization case studies. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):871–880, Jan 2017. doi: 10.1109/TVCG.2016.2599214
- [18] S. E. Fienberg. Graphical methods in statistics. *The American Statistician*, 33(4):165–178, 1979.
- [19] D. L. Gresh. Self-corrected perceptual colormaps. Technical report, IBM, 2008. RC24542 (W0804-104).
- [20] M. Harrower and C. A. Brewer. Colorbrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 2013.
- [21] J. Heer and M. Stone. Color naming models for color selection, image editing and palette design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pp. 1007–1016. ACM, New York, NY, USA, 2012. doi: 10.1145/2207676.2208547
- [22] D. H. Hofstadter. *Gödel, Escher, Bach: An Eternal Golden Braid; [a Metaphoric Fugue on Minds and Machines in the Spirit of Lewis Carroll]*. Penguin Books, 1980.
- [23] R. Huertas, M. Melgosa, and C. Olari. Performance of a color-difference formula based on OSA-UCS space using small-medium color differences. *JOSA A*, 23(9):2077–2084, 2006.
- [24] S. Kaski, J. Venna, and T. Kohonen. Coloring that reveals cluster structures in multivariate data. *Australian Journal of Intelligent Information Processing Systems*, 6(2):82–88, 2000.
- [25] G. Kindlmann, E. Reinhard, and S. Creem. Face-based luminance matching for perceptual colormap generation. In *Proceedings of the conference on Visualization '02*, pp. 299–306. IEEE Computer Society, 2002.
- [26] G. Kindlmann and C. Scheidegger. An algebraic process for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2181–2190, Dec 2014. doi: 10.1109/TVCG.2014.2346325
- [27] F. A. Kingdom. Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, 51(7):652–673, 2011.
- [28] P. Kovési. Good colour maps: How to design them. *arXiv preprint arXiv:1509.03700*, 2015.
- [29] H. Levkowitz. Perceptual steps along color scales. *International Journal of Imaging Systems and Technology*, 7(2):97–101, 1996.
- [30] H. Levkowitz and G. T. Herman. The design and evaluation of color scales for image data. *IEEE Computer Graphics and Applications*, 12(1):72–80, 1992.
- [31] A. Light and P. J. Bartlein. The end of the rainbow? Color schemes for improved data graphics. *Eos*, 85(40):385–391, 2004.
- [32] S. Lin, J. Fortuna, C. Kulkarni, M. Stone, and J. Heer. Selecting semantically-resonant colors for data visualization. In *Proceedings of the 15th Eurographics Conference on Visualization, EuroVis '13*, pp. 401–410. The Eurographics Association & #38; John Wiley & #38; Sons, Ltd., Chichester, UK, 2013. doi: 10.1111/cgf.12127
- [33] M. R. Luo, G. Cui, and C. Li. Uniform colour spaces based on ciecam02 colour appearance model. *Color Research & Application*, 31(4):320–330, 2006. doi: 10.1002/col.20227
- [34] M. R. Luo, G. Cui, and B. Rigg. The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research & Application*, 26(5):340–350, 2001.
- [35] M. R. Luo and C. Li. Ciecama02 and its recent developments. In *Advanced Color Image Processing and Analysis*, pp. 19–58. Springer, 2013.
- [36] D. L. MacAdam. Visual sensitivities to color differences in daylight*. *J. Opt. Soc. Am.*, 32(5):247–274, May 1942. doi: 10.1364/JOSA.32.000247
- [37] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5(2):110–141, Apr. 1986. doi: 10.1145/22949.22950
- [38] M. Mahy, L. Eycken, and A. Oosterlinck. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research & Application*, 19(2):105–121, 1994.
- [39] B. J. Meier, A. M. Spalter, and D. B. Karelitz. Interactive color palette tools. *IEEE Computer Graphics and Applications*, 24(3):64–72, 2004.
- [40] G. W. Meyer and D. P. Greenberg. Perceptual color spaces for computer graphics. *ACM SIGGRAPH Computer Graphics*, 14(3):254–261, 1980.
- [41] S. Mittelstädt, J. Bernard, T. Schreck, M. Steiger, J. Kohlhammer, and D. A. Keim. Revisiting Perceptually Optimized Color Mapping for High-Dimensional Data Analysis. In N. Elmqvist, M. Hlawitschka, and J. Kennedy, eds., *EuroVis - Short Papers*. The Eurographics Association, 2014. doi: 10.2312/eurovisshort.20141163
- [42] S. Mittelstädt, D. Jäckle, F. Stoffel, and D. A. Keim. ColorCAT: Guided Design of Colormaps for Combined Analysis Tasks. In E. Bertini, J. Kennedy, and E. Puppo, eds., *Eurographics Conference on Visualization (EuroVis) - Short Papers*, pp. 115–119. The Eurographics Association, 2015. doi: 10.2312/eurovisshort.20151135
- [43] S. Mittelstädt and D. A. Keim. Efficient contrast effect compensation with personalized perception models. *Computer Graphics Forum*, 34(3):211–220, 2015. doi: 10.1111/cgf.12633
- [44] S. Mittelstädt and D. A. Keim. Efficient Contrast Effect Compensation with Personalized Perception Models. *Computer Graphics Forum*, 2015. doi: 10.1111/cgf.12633
- [45] S. Mittelstädt, A. Stoffel, and D. A. Keim. Methods for compensating contrast effects in information visualization. In *Computer Graphics Forum*, vol. 33, pp. 231–240. Wiley Online Library, 2014.
- [46] K. Moreland. Diverging color maps for scientific visualization. In *International Symposium on Visual Computing*, pp. 92–103. Springer, 2009.
- [47] N. Moroney, M. D. Fairchild, R. W. Hunt, C. Li, M. R. Luo, and T. Newman. The ciecam02 color appearance model. In *Color and Imaging Conference*, vol. 2002, pp. 23–27. Society for Imaging Science and Technology, 2002.
- [48] L. Padilla, P. S. Quinan, M. Meyer, and S. H. Creem-Regehr. Evaluating the impact of binning 2d scalar fields. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):431–440, Jan 2017. doi: 10.1109/TVCG.2016.2599106
- [49] B. Pham. Spline-based color sequences for univariate, bivariate and trivariate mapping. In *Proceedings of the 1st conference on Visualization '90*, pp. 202–208. IEEE Computer Society Press, 1990.
- [50] S. H. Pizer, J. B. Zimmerman, and R. E. Johnston. Concepts of the display

- of medical images. *IEEE Transactions on Nuclear Science*, 29(4):1322–1330, 1982.
- [51] S. M. Pizer. Intensity mappings to linearize display devices. *Computer Graphics and Image Processing*, 17(3):262–268, 1981.
- [52] S. M. Pizer, J. B. Zimmerman, and R. E. Johnston. Contrast transmission in medical image display. In *1st International Symposium on Medical Imaging and Image Interpretation*, pp. 2–9. International Society for Optics and Photonics, 1982.
- [53] D. Raj Pant and I. Farup. Riemannian formulation and comparison of color difference formulas. *Color Research & Application*, 37(6):429–440, 2012.
- [54] H. L. Resnikoff. Differential geometry and color perception. *Journal of Mathematical Biology*, 1(2):97–131, 1974.
- [55] P. L. Rheingans. Task-based color scale design. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making*, pp. 35–43. International Society for Optics and Photonics, 2000.
- [56] P. K. Robertson. A methodology for scientific data visualisation: choosing representations based on a natural scene paradigm. In *Proceedings of the 1st conference on Visualization'90*, pp. 114–123. IEEE Computer Society Press, 1990.
- [57] P. K. Robertson and J. F. O'Callaghan. The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications*, 6(2):24–32, 1986.
- [58] B. E. Rogowitz and A. D. Kalvin. The "Which Blair Project": a quick visual method for evaluating perceptual color maps. In *Visualization, 2001. VIS'01. Proceedings*, pp. 183–556. IEEE, 2001.
- [59] B. E. Rogowitz, A. D. Kalvin, A. Pelah, and A. Cohen. Which trajectories through which perceptually uniform color spaces produce appropriate colors scales for interval data? In *7th Color and Imaging Conference*, pp. 321–326. Society for Imaging Science and Technology, 1999.
- [60] B. E. Rogowitz and L. A. Treinish. Data visualization: the end of the rainbow. *IEEE spectrum*, 35(12):52–59, 1998.
- [61] B. E. Rogowitz, L. A. Treinish, and S. Bryson. How not to lie with visualization. *Computers in Physics*, 10(3):268–273, 1996.
- [62] P. O. Runge and H. Steffens. *Farben-Kugel oder Construction des Verhältnisses aller Mischungen der Farben zu einander, ... Nebst einer Abh. über die Farben in der Natur, von Henrik Steffens.-Hamburg, Perthes 1810. 60 S., 2 Tf.* Perthes, 1810.
- [63] J. W. S ammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on computers*, 18(5):401–409, 1969.
- [64] F. Samsel, T. L. Turton, P. Wolfram, and R. Bujack. Intuitive Colormaps for Environmental Visualization. In K. Rink, A. Middel, D. Zeckzer, and R. Bujack, eds., *Workshop on Visualisation in Environmental Sciences (EnvirVis)*. The Eurographics Association, 2017. doi: 10.2312/envirvis.20171105
- [65] K. Schloss, C. Gramazio, and C. Walmsley. Which color means more? an investigation of color-quantity mapping in data visualization. *Journal of vision*, 15(12):1317–1317, 2015.
- [66] K. B. Schloss and S. E. Palmer. Aesthetics of color combinations. vol. 7527, pp. 752712–752731. doi: 10.1117/12.849111
- [67] P. Schulze-wollgast, C. Tominski, and H. Schumann. Enhancing visual exploration by appropriate color coding. In *Proceedings of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, pp. 203–210, 2005.
- [68] S. Silva, J. Madeira, and B. S. Santos. There is more to color scales than meets the eye: a review on the use of color in visualization. In *Information Visualization, 2007. IV'07. 11th International Conference*, pp. 943–950. IEEE, 2007.
- [69] S. Silva, B. S. Santos, and J. Madeira. Using color in visualization: A survey. *Computers & Graphics*, 35(2):320–333, 2011.
- [70] N. Silvestrini and E. Fischer. Idee farbe. *Farbsysteme in Kunst und Wissenschaft. Baumann & Stroemer, Zürich*, 1994.
- [71] K. R. Sloan and C. M. Brown. Color map techniques. *Computer Graphics and Image Processing*, 10(4):297–317, 1979.
- [72] I. Spence, N. Kutlesa, and D. L. Rose. Using color to code quantity in spatial displays. *Journal of Experimental Psychology: Applied*, 5(4):393, 1999.
- [73] M. Stone, D. A. Szafir, and V. Setlur. An engineering model for color difference as a function of size. In *22nd Color and Imaging Conference Final Program and Proceedings*, 22nd Color and Imaging Conference, pp. 253–258. Society for Imaging Science and Technology, 2014.
- [74] J. Tajima. Uniform color scale applications to computer graphics. *Computer Vision, Graphics, and Image Processing*, 21(3):305–325, 1983.
- [75] R. Tamassia. *Handbook of Graph Drawing and Visualization*. Discrete Mathematics and Its Applications. Taylor & Francis, 2013.
- [76] D. C. Thompson, J. Bennett, C. Seshadhri, and A. Pinar. A provably-robust sampling method for generating colormaps of large data. In *LDAV*, pp. 77–84, 2013.
- [77] K. M. Thyng, C. A. Greene, R. D. Hetland, H. M. Zimmerle, and S. F. DiMarco. True colors of oceanography. *Oceanography*, 29(3):10, 2016.
- [78] C. Tominski, G. Fuchs, and H. Schumann. Task-driven color coding. In *2008 12th International Conference Information Visualisation*, pp. 373–380. IEEE, 2008.
- [79] B. E. Trumbo. A theory for coloring bivariate statistical maps. *The American Statistician*, 35(4):220–226, 1981.
- [80] S. van der Walt and N. Smith. Matplotlib documentation update. bids.github.io/colormap/.
- [81] H. Wainer and C. M. Francolini. An empirical inquiry concerning human understanding of two-variable color maps. *The American Statistician*, 34(2):81–93, 1980.
- [82] N. Waldin, M. Bernhard, P. Rautek, and I. Viola. Personalized 2D color maps. *Computers & Graphics*, 59:143–150, 2016.
- [83] C. Ware. Color sequences for univariate maps: Theory, experiments and principles. *IEEE Computer Graphics and Applications*, 8(5):41–49, 1988.
- [84] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3 ed., 2012.
- [85] C. Ware, T. L. Turton, F. Samsel, R. Bujack, and D. H. Rogers. Evaluating the Perceptual Uniformity of Color Sequences for Feature Discrimination. In K. Lawonn, N. Smit, and D. Cunningham, eds., *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3)*. The Eurographics Association, 2017. doi: 10.2312/eurorv3.20171107
- [86] M. Wijffelaars, R. Vliegen, J. J. Van Wijk, and E.-J. Van Der Linden. Generating color palettes using intuitive parameters. *Computer Graphics Forum*, 27(3):743–750, 2008. doi: 10.1111/j.1467-8659.2008.01203.x
- [87] W. D. Wright and F. H. G. Pitt. Hue-discrimination in normal colour-vision. *Proceedings of the Physical Society*, 46(3):459, 1934.
- [88] A. Zeileis, K. Hornik, and P. Murrell. Escaping RGBland: selecting colors for statistical graphics. *Computational Statistics & Data Analysis*, 53(9):3259–3270, 2009.
- [89] H. Zhang and E. Montag. Perceptual color scales for univariate and bivariate data display. *The Society for Imaging Science and Technology (IS&T)*, 2006.
- [90] L. Zhou and C. Hansen. A survey of colormaps in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(8):2051–2069, 2016.