

# Dynamic Color Assignment for Hierarchical Data

Jiashu Chen\* , Weikai Yang\* , Zelin Jia, Lanxi Xiao, and Shixia Liu 

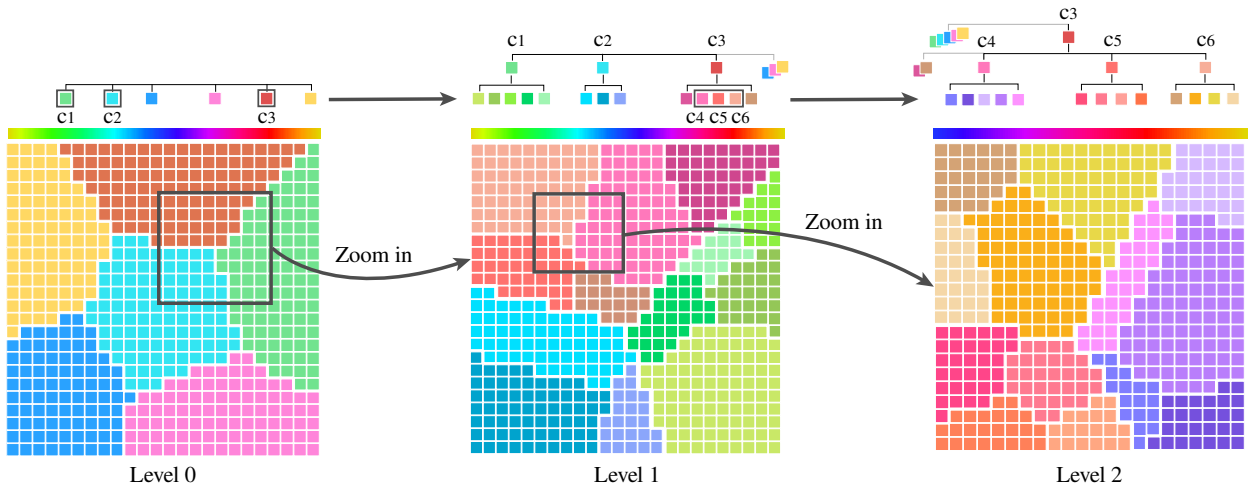


Fig. 1: Based on user exploration, our method dynamically selects the color range and assigns colors to classes within the range, which ensures high discriminability and harmony at each level and maintains consistency across different levels.

**Abstract**—Assigning discriminable and harmonic colors to samples according to their class labels and spatial distribution can generate attractive visualizations and facilitate data exploration. However, as the number of classes increases, it is challenging to generate a high-quality color assignment result that accommodates all classes simultaneously. A practical solution is to organize classes into a hierarchy and then dynamically assign colors during exploration. However, existing color assignment methods fall short in generating high-quality color assignment results and dynamically aligning them with hierarchical structures. To address this issue, we develop a dynamic color assignment method for hierarchical data, which is formulated as a multi-objective optimization problem. This method simultaneously considers color discriminability, color harmony, and spatial distribution at each hierarchical level. By using the colors of parent classes to guide the color assignment of their child classes, our method further promotes both consistency and clarity across hierarchical levels. We demonstrate the effectiveness of our method in generating dynamic color assignment results with quantitative experiments and a user study.

**Index Terms**—Color assignment, Hierarchical Visualization, Discriminability, Harmony.

## 1 INTRODUCTION

Assigning colors to samples according to their class labels and spatial distribution is a common practice in data analysis [35, 70, 72]. A high-quality color assignment should be discriminable and harmonic to ensure clarity and attractiveness in visualizations [20, 34]. However, as the number of classes increases, it is challenging to select a large number of colors that are easily distinguishable from each other but also harmonize together [17, 56]. A practical solution is to organize classes into a hierarchical structure and then dynamically assign consistent colors across hierarchical levels. This solution not only enhances scalability by reducing the requirement for a large number of distinct colors in a static visualization, but also alleviates cognitive load during data exploration. However, existing color assignment methods do not fully support this solution. Some methods generate color assignment results independently for each hierarchical level, which cannot maintain color

consistency across levels [17, 36]. This will disrupt the user’s mental map during exploration. Other methods apply strict color constraints to achieve consistency but sacrifice discriminability and harmony at each level [19, 55]. These limitations highlight the need for a dynamic color assignment method that is capable of maintaining color discriminability and harmony at each level and ensures consistency across different levels during exploration.

To determine the design requirements for developing such a dynamic color assignment method, we first conduct interviews with six experts specializing in Information Design in a School of Arts. The findings indicate that the most important goal is to ensure discriminability, which facilitates identifying class labels of data samples. Following this, harmony is identified as the second most important factor, critical to producing visually attractive results that engage users. They also point out that considering spatial distribution can further improve discriminability and harmony and thus facilitate data analysis. For example, class boundaries can be made clearer by increasing discriminability between adjacent classes [36]. Accordingly, we formulate the color assignment as a multi-objective optimization problem with suggested priorities among these objectives [39], and apply the most advanced theories to quantify each objective. Next, to solve this complex optimization problem, we employ simulated annealing for its high flexibility in handling multiple objectives, and combine it with the continuation method to sequentially incorporate discriminability, harmony, and spatial distribution during the optimization process. This

\* J. Chen, W. Yang, Z. Jia, L. Xiao, S. Liu are with the School of Software, BNRist, Tsinghua University. J. Chen and W. Yang are joint first authors. S. Liu is the corresponding author. E-mail: {{cjs22,yangwk21,zlin,lanxi}@mails., shixia@} tsinghua.edu.cn

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxxx

accelerates convergence to a better solution by guiding the optimization process towards more promising regions in the solution space [2]. To generate dynamic color assignment results based on user exploration and align them with the hierarchical structures within datasets, the colors of the parent classes are used to guide the color assignment of their child classes. This is achieved by dynamically selecting appropriate color ranges for child classes based on the colors of their parent classes, and then optimizing the color assignment result within the selected color range. As shown in Fig. 1, our method achieves color consistency across levels in hierarchical grid visualizations.

Quantitative experiments show that compared to state-of-the-art methods, our method performs best in ensuring discriminability and aligning with hierarchical structures, while still offering comparable levels of harmony. A user study with 20 experts further confirms that our method generates high-quality color assignment results that are closely aligned with user preference.

The main contributions of our work include:

- A color assignment method that achieves better discriminability and harmony.
- A dynamic color range selection method in which the colors of the parent classes guide the color assignment of their child classes.
- An open-source implementation of the proposed color assignment method in both C++ and JavaScript, available at <https://github.com/thu-vis/Dynamic-Color>.

## 2 RELATED WORK

Existing color assignment methods can be classified into two categories based on how they organize classes in a dataset: flat color assignment and hierarchical color assignment.

**Flat color assignment** methods assign colors to all classes without considering their hierarchical relationships. In flat color assignment, ensuring discriminability between colors of different classes is a fundamental requirement and consistently draws research attention over the years [22, 41, 43, 57, 76]. As a pioneering study, Healey [22] proposed a rule-based method to select discriminable colors on the hue wheel that maximize perceptual differences and name differences. Later studies improve discriminability by incorporating more advanced theory in quantifying perceptual differences and/or name differences [17, 17, 51]. For example, Fang *et al.* [17] calculated perceptual differences using CIEDE2000 [52], which improves perceptual uniformity and achieves better alignment with human perception. Setlur *et al.* [51] utilized the name distance proposed by Heer and Stone [23], which includes 153 popular color names and their color-name associations. This offers a more precise way to measure name differences.

In addition to color discriminability, it is also important to generate harmonic and visually appealing color assignment results [11, 20, 25, 31, 34, 53, 61, 63, 75]. For example, Cohen-Or *et al.* [11] introduced a color harmonization method by aligning colors with Matsuda’s established harmonic templates [40]. Color Crafting [53] summarizes the templates of designer-crafted color assignment results in the color space and then generates more color assignment results that mimic designer practices.

Color discriminability and harmony can be further improved by considering the spatial distribution of the visualized data [10, 29, 36, 37, 62]. Wang *et al.* [62] considered the color discriminability between neighboring points and their contrasts to the background in a scatterplot. They then employed a genetic algorithm to find the best color assignment result from a set of pre-defined colors that maximizes discriminability.

Palettailor [36] advances this method by simultaneously adjusting and assigning colors during the optimization process.

Although recent efforts in flat color assignment have achieved certain success in producing discriminable and harmonic colors, they struggle in real-world scenarios where the number of classes can reach hundreds or even thousands. To address the scalability issue in flat color assignment methods, researchers have developed several **hierarchical color assignment** methods. Early efforts focus on generating coherent colors for a static visualization, where all colors are visible at the same time [15, 19, 55]. For example, Fua *et al.* [19] proposed a proximity-based coloring method for hierarchical parallel coordinates. It recursively assigns colors to child classes within a range centered on the color of their parent class. This range becomes progressively narrower at each level, which ensures that the colors of classes of the same parent are more similar to each other than those of different parents. Similarly, Tree Colors [55] divides the hue wheel into several ranges and assigns each to different branches of a tree. As the level increases, the hue range for each branch narrows down, while the saturation increases. This method generates a diverse but consistent color assignment result across the hierarchy. However, as the number of classes increases exponentially with the levels, these methods still suffer from scalability issues regarding color discriminability. To address this issue, later efforts adopt dynamic color assignment that only assigns colors to visible data during exploration [59, 60]. These methods better exploit the color space and thus improve the overall quality of color assignment results. For example, Chameleon [59] uses a force-based method to dynamically adjust color ranges on the hue wheel. This method aims to keep the ranges close to their original position while reducing overlaps between adjacent ranges. However, the force-based method does not prevent color overlaps between child classes of different parents. Cuttlefish [60] extends it to eliminate overlaps by imposing hard constraints. It allows a larger shift on the hue wheel to ensure distinct color ranges.

While these hierarchical color assignment methods succeed in maintaining color consistency between parent classes and child classes, they still face two issues. First, these methods do not simultaneously consider discriminability, harmony, and spatial distribution to produce a high-quality color assignment result at each level. Second, the strict constraints between the colors of parent classes and child classes often lead to insufficient discriminability, especially among the child classes of the same parent. In comparison, we achieve a well-balanced integration of discriminability, harmony, and spatial distribution in our optimization process by formulating color assignment as a multi-objective optimization problem with suggested priorities among these tasks. We also developed an improved color range selection method to enhance discriminability without sacrificing color consistency across class hierarchies. The detailed comparison between representative color assignment methods and our method is summarized in Table 1.

## 3 REQUIREMENT ANALYSIS

We worked closely with six experts (E1-E6) during the development of the dynamic color assignment method. All of them major in Information Design in a School of Arts and have more than 5 years of experience in designing colors for visualizations and/or user interfaces. E1 is the co-author of this paper, while E2-E6 are not. We conducted six semi-structured interviews with each expert to collect the requirements for dynamic color assignment. Initially, we shared the results of existing color assignment methods (*e.g.*, Fig. 2) with the experts. Then,

Table 1: The comparison between several representative color assignment methods and our method.

Methods	Discrimination	Harmony	Spatial Distribution	Alignment with Hierarchy	Dynamic Assignment
Palettailor [36]	✓	-	✓	-	-
Color Crafting [53]	✓	✓	-	-	-
Tree Colors [55]	✓	-	-	✓	-
Cuttlefish [60]	✓	-	-	✓	✓
Ours	✓	✓	✓	✓	✓

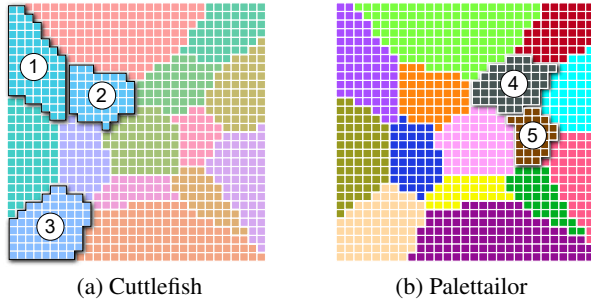


Fig. 2: Results generated by existing color assignment methods.

they were asked to evaluate these results, highlighting both strengths and weaknesses. They were also recommended to modify the colors to express their preferences. Finally, we collected their advice on generating high-quality color assignment results and the factors that warrant particular attention. Each interview lasted between 35 and 45 minutes. In addition to these interviews, we also engaged in biweekly free-form discussions to showcase our color assignment results and promptly collect their feedback.

Based on the six semi-structured interviews, the biweekly discussions, and the literature review, we summarized four design requirements for dynamic color assignment.

**Ensure color discriminability.** All the experts agreed that color discriminability is the most important factor and should be considered first. E1 commented that a minimum threshold of color differences is required to quickly identify different class labels. This is also reflected in several previous research [5, 20, 36, 54]. When examining color assignment results generated by existing methods, E2 and E4 pointed out that Color Crafting [53] and Cuttlefish [60] failed to achieve good color discriminability when the number of colors exceeded 10. As shown in Fig. 2(a), the color assignment result generated by Cuttlefish results in three colors  $\blacksquare\blacksquare\blacksquare$  that are not sufficiently distinguishable from each other. This is because this method mainly considers the difference in the hue channel, which limits its selection range. A better color discriminability can be achieved by modifying their saturation and luminance ( $\blacksquare\blacksquare\blacksquare$  vs.  $\blacksquare\blacksquare\blacksquare$ ). Therefore, it is necessary to simultaneously consider differences in hue, saturation, and luminance when generating color assignment results.

**Enhance color harmony.** All the experts pointed out that they would also consider color harmony when choosing colors in their designs. In this process, they would avoid using strongly disliked colors, such as DarkSlateGray ( $\blacksquare$ ) and SaddleBrown ( $\blacksquare$ ) in Fig. 2(b). The strategy of excluding strongly disliked color ranges has also been widely employed in existing color assignment methods [20, 36]. When discussing how to improve color harmony, three experts pointed out that according to harmonic template theory [11, 40], harmonic colors usually conform to specific geometric patterns in the color space, such as the hue wheel. Four experts also noted that, in addition to the hue wheel, the balance between saturation and luminance is also crucial to color harmony.

**Consider spatial distribution.** Our experts also highlighted the importance of considering spatial distribution in generating color assignment results. Taking spatial distribution into account not only enhances data analysis [29, 62] but also improves the aesthetic appeal of the results [24, 30, 46]. E3 and E6 emphasized that the assigned colors of two spatially adjacent classes significantly affect perception and thus deserve careful consideration. For example, enhancing the color discriminability between spatially adjacent classes can make the class boundaries clearer and aid in identifying different classes. Moreover, maintaining color harmony between adjacent classes can produce more visually pleasing results. E5 also noted that assigning similar colors to similar classes would facilitate data understanding and exploration.

**Align with hierarchical structures.** As rich hierarchies are ubiquitous in datasets [6, 71], four experts also acknowledged that the generated color assignment results should accurately reflect the hierarchical structures. E1 said, “It is common practice to use similar colors to encode a

parent class and its corresponding child classes, which facilitates the identification of hierarchical relationships and keeps the user’s mental map throughout the zooming process.” E4 further emphasized that to avoid misinterpretation of hierarchical relationships, the color differences between child classes of the same parent class should be smaller than those between child classes of different parent classes.

## 4 DYNAMIC COLOR ASSIGNMENT

### 4.1 Method Overview

Driven by the identified requirements, we propose a dynamic color assignment method that aligns well with the class hierarchy across levels. At each level, our method simultaneously considers discriminability, harmony, and spatial distribution. As shown in Fig. 3, our method consists of two modules: **color range selection** and **color assignment**. The color range selection module selects an appropriate color range to ensure consistency across hierarchical levels, and the color assignment module generates high-quality color assignment results within the selected color range. Specifically, when assigning colors for top-level classes, the color range selection module selects the full color range that allows greater flexibility for generating color assignment results. When users focus on a specific region for closer examination, this module selects the appropriate color range for the child classes based on the colors of their parent classes. This ensures color consistency and provides a coherent exploration experience that adapts to user interactions. Based on this exploration process, we will first introduce how to generate high-quality color assignment results at each level and then describe how to ensure consistency across levels using dynamic color range selection.

### 4.2 Color Assignment

The expert interviews reveal that generating high-quality color assignment results requires optimizing multiple goals, including discriminability, harmony, and spatial distribution. However, due to the conflicting nature of these goals, it is impossible to maximize all of them simultaneously. For example, the color assignment result with the best discriminability will contain colors with extremely high luminance, leading to lower harmony. In practice, there are multiple Pareto-optimal solutions that cannot be enhanced in one goal without compromising another. Existing efforts require users to adjust the weighting parameters of different goals to explore different Pareto-optimal solutions. However, it brings an extra burden for users to fine-tune the weighting parameters and determine a better one. In the requirement analysis, we have identified a priority order of *discriminability* > *harmony* > *spatial distribution*. Therefore, we utilized a priority-specific Pareto-optimal strategy, which is effective in identifying the Pareto-optimal solution that satisfies the priority order [39]. This is achieved by guaranteeing that goals with higher priorities have higher objective values. Specifically, given a set of colors  $c_1, c_2, \dots, c_m$ , the optimization problem is formulated as:

$$\begin{aligned} & \max_{c_1, c_2, \dots, c_m} E_D + \alpha E_H + \beta E_{SD}, \\ & \text{s.t. } E_D \geq E_H \geq E_{SD}; \quad c_i \in \mathcal{C}, \forall i \in \{1, 2, \dots, m\}. \end{aligned} \quad (1)$$

Here,  $E_D$ ,  $E_H$ , and  $E_{SD}$  represent the objective values for discriminability, harmony, and spatial distribution, respectively.  $\mathcal{C}$  is the feasible color range. The weighting parameters  $\alpha$  and  $\beta$  control the trade-offs between multiple objectives, which will be automatically determined during the optimization process.

#### 4.2.1 Color Discriminability

Following Palettaylor [36], the total objective function of color discriminability  $E_D$  consists of two terms: perceptual difference  $E_{PD}$  and name difference  $E_{ND}$ .

**Perceptual difference.** Perceptual difference quantifies the human-perceived difference between two colors. In our implementation, we use the CIEDE2000 formula to calculate this perceptual difference because it is closely aligned with human perception [52]. Accordingly, the perceptual difference is defined as:

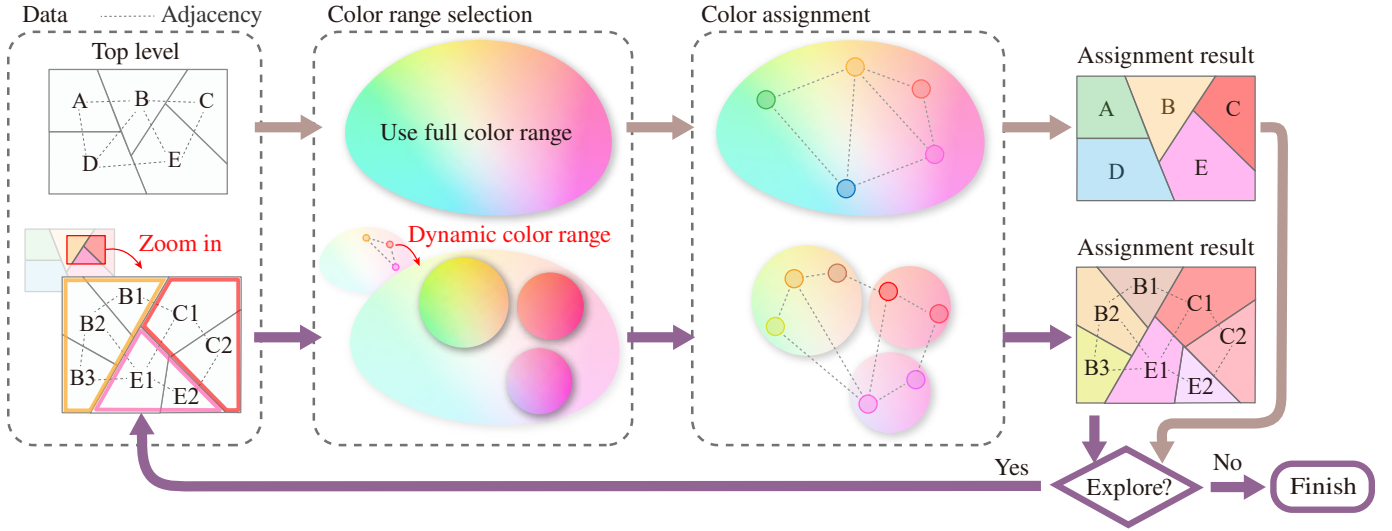


Fig. 3: Method overview. Given data, the color range selection module selects an appropriate range of colors. Then, the color assignment module generates a discriminable and harmonic color assignment result within that range.

$$E_{PD} = \min_{1 \leq i < j \leq n} D(c_i, c_j) + \min(\min_{1 \leq i < j \leq n} D(c_i, c_j) - 10, 0), \quad (2)$$

where  $D(c_i, c_j)$  is the perceptual difference between colors  $c_i$  and  $c_j$  using the CIEDE2000 formula. The first term aims to maximize the minimal perceptual difference among all color pairs. The second term introduces an extra penalty when the minimal perceptual difference falls below a threshold of 10, which is required to achieve high accuracy in judging whether two colors are identical.

**Name difference.** In practice, colors that are perceptually different may still be described using the same name. For example, these two colors ■ and ■ are both commonly described as “Blue.” Such naming ambiguity should be avoided since it leads to confusion when discussing colors in visualizations. Heer and Stone [23] introduced the concept of name difference to quantify the likelihood that two colors are described using the same name. They represented each color with a 153-dimensional feature vector, where each dimension corresponds to a popular color name. Name difference is calculated by averaging the cosine distances between all color pairs:

$$E_{ND} = \frac{2}{m(m-1)} \sum_{1 \leq i < j \leq n} (1 - \cos \frac{T_{c_i} \cdot T_{c_j}}{\|T_{c_i}\| \cdot \|T_{c_j}\|}), \quad (3)$$

where  $T_{c_i}$  and  $T_{c_j}$  are the feature vectors of colors  $c_i$  and  $c_j$ .

The color discriminability is then defined as a weighted combination of perceptual difference and name difference:  $E_D = \gamma_1 E_{PD} + \gamma_2 E_{ND}$ . Following Palettailor [36], we set  $\gamma_1 = 0.1$  and  $\gamma_2 = 2.0$ , which gives satisfactory results in practice.

#### 4.2.2 Color Harmony

We adopt the most advanced color harmony theory developed by Lara-Alvarez and Reyes [27]. We choose this theory because it aligns well with human preferences and uses hue-chroma-lightness (CIELCh) color space, which is more uniform and thus suitable for research on optimizing colors. According to this theory, a harmonic color assignment result should follow specific patterns on both the hue wheel and the chroma-lightness plane. Accordingly, the total objective function of color harmony  $E_H$  consists of two terms: hue harmony  $E_{Hue}$  and chroma-lightness harmony  $E_{CL}$ .

**Hue harmony.** We do not directly use the hue harmony term proposed by Lara-Alvarez and Reyes [27] because it oversimplifies the widely-used Matsuda’s harmonic templates [40] defined on the hue-saturation-value (HSV) color space and directly applies them to the hue wheel

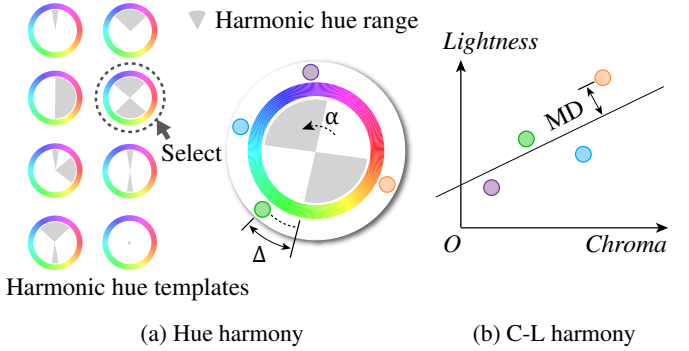


Fig. 4: Color harmony: (a) on the hue wheel, colors that lie in the gray region are considered harmonic; (b) on the chroma-lightness plane, colors that follow a straight line are considered harmonic.

of the CIELCh space. The large difference between the hue wheels of these two spaces often leads to large discrepancies in color perception and harmony. To tackle this issue, we use Matsuda’s templates to measure the hue harmony. Given a color assignment result, we first compute the corresponding hue values in the HSV space and then compare them with Matsuda’s templates. The difference between a color assignment result and a hue template is quantified by summing the minimal angular distances between the hue value of each color in the color assignment result and the hue range of the template, which can be freely rotated. Specifically, let  $\theta_1, \dots, \theta_m$  be the hue values of colors  $c_1, \dots, c_m$  in a color assignment result  $C$ , and  $R$  be the hue range of a harmonic hue template, the hue difference between them is:

$$\text{Hue\_Diff}(C, R) = \min_{\alpha \in [0, 360^\circ]} \sum_{i=1}^m \Delta(\alpha + \theta_i \bmod 360^\circ, R), \quad (4)$$

where  $\alpha$  is the rotation angle of the hue range, and  $\Delta(\theta, R)$  is the minimal angular distance between the value  $\theta$  and the range  $R$ . If  $\theta$  falls within the range  $R$ , the minimal angular distance is 0. Otherwise, it is calculated as the distance between  $\theta$  and the nearest boundary of  $R$ . The objective value of hue harmony is then defined as the negative value of the smallest hue difference across all eight hue templates, and it is normalized to the range  $[0, 1]$  using min-max normalization.

$$E_{Hue} = \text{Normalize}(-\min_R \text{Hue\_Diff}(C, R)). \quad (5)$$

**Chroma-lightness harmony.** We directly used the chroma-lightness harmony term proposed by Lara-Alvarez and Reyes [27], which encourages colors to follow a straight line in the chroma-lightness plane. As shown in Fig. 4(b), given a color assignment result, it first determines the corresponding maximum likelihood line that best fits the colors in the chroma-lightness plane. Next, it calculates the deviation of  $i$ -th color from this ideal line, denoted by  $MD_i$ . As it is unnecessary to strictly adhere to the line, we allow a deviation of 15 units on the chroma-lightness plane, which is recommended by Liu *et al.* [34]. The optimization objective is then defined as:

$$E_{LC} = \text{Normalize}(-\sum_{i=1}^n \max(MD_i - 15, 0)). \quad (6)$$

The color harmony is then defined as the sum of hue harmony and chroma-lightness harmony, which gives satisfactory results in practice.

### 4.2.3 Spatial Distribution

When applying color assignment results in visualization, discriminability and harmony can be further enhanced by considering the spatial distribution of the visualized data. For example, increasing the color differences between adjacent classes can enhance the color discriminability and make the boundaries clearer [29, 36], and ensuring the harmony of color pairs between adjacent classes can generate more visually coherent results [46, 47]. Since the concept of adjacent classes changes with different types of visualizations, the optimization objective for data distribution is calculated by averaging the score of all neighboring sample pairs:

$$E_{SD} = \frac{1}{|X|} \sum_{x_i \in X} \frac{1}{|\Omega_{x_i}|} \sum_{x_j \in \Omega_{x_i}} \frac{f(x_i, x_j)}{d(x_i, x_j)}. \quad (7)$$

Here,  $X$  is the set of all samples, and  $\Omega_{x_i}$  is the set of neighboring samples of sample  $x_i$  in visualization. We consider three representative types of visualization: scatterplots [69] (point-based), parallel coordinates [36] (line-based), and grid visualizations [7, 32] (area-based). For grid visualizations, we consider eight surrounding cells of the center cell. For scatterplots and parallel coordinates, we use eight nearest neighbors, which is consistent with grid visualizations. The score of each sample pair is  $f(x_i, x_j)/d(x_i, x_j)$ . Here,  $d(x_i, x_j)$  is the spatial distance between samples  $x_i$  and  $x_j$  in the visualization. Therefore, a closer pair would have more impact on the objective.  $f(x_i, x_j)$  is the optimization objective for each pair. Based on the requirement analysis, we consider two modes for this term: difference mode and similarity mode.

The difference mode favors a color pair with a larger perceptual difference between adjacent classes. Therefore,  $f(x_i, x_j)$  is set as  $D(c(x_i), c(x_j)) + P(c(x_i), c(x_j))$ . The first term is the perceptual difference between the colors of  $x_i$  and  $x_j$ . The second term measures the pair harmony between colors of  $x_i$  and  $x_j$ . We use the formula introduced in Ou's recent work about color pair harmony [48] to calculate this term.

The similarity mode reduces the perceptual difference between similar classes to facilitate data exploration and understanding. In our implementation,  $f(x_i, x_j)$  is defined as  $-D(c(x_i), c(x_j)) \cdot s(x_i, x_j) + P(c(x_i), c(x_j))$ , where  $s$  measures the class similarity. This similarity is determined by first averaging the feature vectors within each class to create class-level feature vectors and then calculating the similarities between them. In the similarity mode, a larger perceptual difference for more similar class pairs results in a higher penalty.

### 4.2.4 Optimization





A straightforward way to solve the optimization problem defined in Eq. (1) is simulated annealing. It is chosen because of its flexibility to accommodate multiple objectives and its effectiveness in escaping local optima during the optimization process. First, we use the blue noise sampling [12] to generate the initial color assignment result within the default color range. This technique ensures that colors are evenly distributed, which provides basic discriminability. At each iteration, the

algorithm adjusts the color assignment result and re-evaluates the objective value. Adjustments that improve the objective value are always accepted, while those that reduce the objective value are accepted with a progressively decreasing probability over time. However, the simulated annealing algorithm suffers from slow convergence due to the low acceptance rate that comes with the original highly non-convex problem. To accelerate convergence, we combine it with the continuation method. Starting from solely optimizing discriminability, it sequentially incorporates harmony and spatial distribution. Each solution to the previous problem serves as a starting point for the optimization of the subsequent problem. By guiding the optimization process towards more promising regions in the solution space, this method achieves higher acceptance rates and faster convergence [2]. The experimental result shows that our method can generate color assignment results for 30 classes in 1 second, which well supports real-time interaction for users navigating through hierarchical visualizations (see the supplemental material for more details). When incorporating a new optimization goal in each stage of the continuation method, we dynamically set its weighting parameter using loss-balanced task weighting [33, 68]. The basic idea is to ensure that different goals are optimized in a similar progress. Thus, the weighting parameter of the goal with less progress will be increased so that it can be further improved in subsequent iterations. Specifically, in each stage, the weighting parameter of the newly incorporated goal is set as the ratio between the current objective value and the possibly maximal value. After the algorithm converges, the weighting parameter will be fixed, and the continuation method will move to the next stage. Fig. 5 shows the incremental refinement of the color assignment results through each phase of the continuation method. The idea of the continuation method also aligns with the typical process of hand-crafted color assignment design. Initially, users select a set of distinguishable colors tailored to the number of classes. Next, they adjust the hue, chroma, and lightness to improve harmony. Finally, they assign these colors to different classes in the visualization and make slight adjustments to refine the overall visual effects.

## 4.3 Color Range Selection

When generating the color assignment result for classes at the top level, we use the default color range during the optimization process. When generating the color assignment result for classes at the deeper level, we select an appropriate color range based on parent classes selected by users. The colors of their child classes will be restricted within the selected color range. This ensures that the generated color assignment result reflects the hierarchical structures within the data.

### 4.3.1 Default Color Range

Guided by expert interviews and the common practice in color assignment research [13, 20, 55, 60], we set the default range for chroma and lightness as [40, 85] instead of the full range [0, 100]. This excludes relatively extreme colors, including dim colors with low chroma/lightness (e.g., ) or highly intense colors that are glaring (e.g., ). Moreover, some studies have pointed out that even within this range, there are still some strongly disliked colors [49, 73], such as  and . To address this issue, we further exclude the range where lightness falls within [40, 75] and hue simultaneously falls within [85°, 114°], as previously utilized by Gramazio *et al.* [20].

### 4.3.2 Dynamic Color Range

For each parent class, we dynamically select a sphere based on the perceptual difference to guide the color assignment of its child classes. We use spheres here because they allow us to ensure that the colors under the same parent are more similar than the colors under different parents by determining proper radii of spheres. However, when the colors of parent classes are close to each other, the available color range of each sphere becomes too narrow to generate high-quality color assignment results. This issue becomes more common and more severe when users explore deeper hierarchical levels. In practice, reusing the color range of invisible classes after zooming in can better exploit the color space and will not cause much confusion [59, 60]. As shown

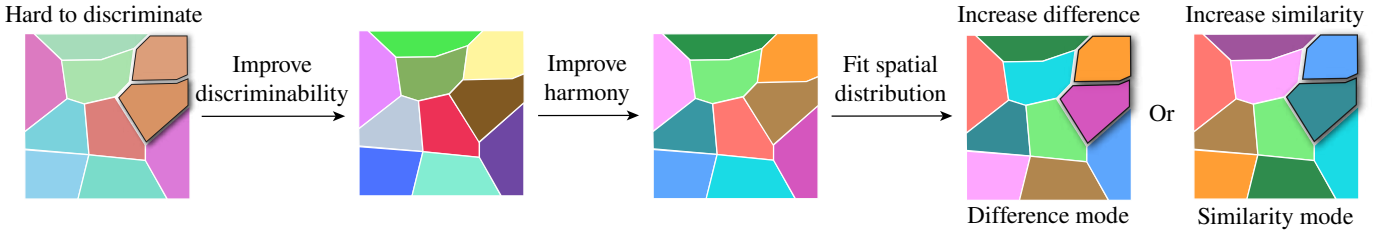


Fig. 5: The continuation method improves assignment result by sequentially incorporating discriminability, harmony, and spatial distribution.

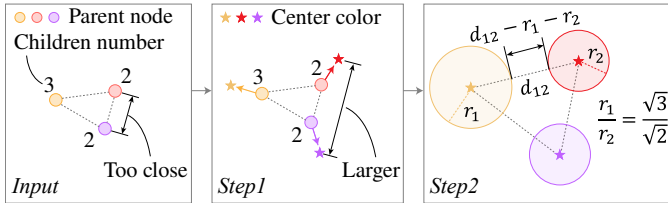


Fig. 6: The workflow of dynamic color range selection.

in Fig. 6, this can be achieved by first adjusting the centers of these spheres and then determining their radii.

**Step 1: Adjust the centers of the spheres.** We use the color assignment method introduced in Sec. 4.2 to adjust the centers of the spheres based on the colors of parent classes. We choose it because of its effectiveness in ensuring discriminability and harmony. However, directly applying this method can lead to two issues. First, without a proper constraint on the adjustment range, colors may change excessively and thus increase the recognition burden. Second, if the chroma or lightness of the colors of parent classes are close to the boundary of the default range ([40, 85]), it leaves less space to create aesthetically pleasing colors for child classes. To address these two issues, we impose additional constraints during the simulated annealing process to adjust the sphere centers. First, we ensure consistency between the initial colors and the adjusted colors. Specifically, each adjusted color must remain closest to its initial color. By doing so, users can better maintain their mental map and correctly connect the initial colors with the adjusted colors. Second, we narrow down the feasible range for chroma and lightness from [40, 85] to [45, 80]. This provides more opportunities to generate high-quality color assignment results for child classes.

**Step 2: Determine the radii of the spheres.** The goal of determining the radii is to ensure that the colors within the same sphere are closer than colors in different spheres and that the spheres with more child classes have larger radii to maintain discriminability. Let  $r_i$  and  $r_j$  denote the radii of two spheres, and  $d_{ij}$  denotes the distance between their centers. The gap between these two spheres will be  $d_{ij} - r_i - r_j$ .

First, we ensure that the gap must exceed the radii of both spheres, *i.e.*,  $d_{ij} - r_i - r_j > \max(r_i, r_j)$ . However, during the development of our methods, the experts pointed out that the hue plays a more important role in the identification of parent-child relationships, and only considering perceptual differences can sometimes result in misunderstanding. For example, in the left image, the colors are

sampled within the ranges of three spheres: green (top), blue (bottom-left), and orange (bottom-right). Although the color on the top-right corner (■) is closer to the green color than the orange and the blue, there is still a noticeable difference in its hue compared to the green color. This potentially misleads users to perceive it as belonging to a separate parent class, *e.g.*, a parent class with yellow color. To avoid such misunderstanding, we add an additional restriction on hue in a similar way, which ensures that the gap between two hue ranges must exceed the length of both ranges. Second, we study how to properly determine the radius of the sphere based on the number of child classes. We conduct an experiment to estimate the relationship between the radius and the number of child classes through blue noise sampling [12, 74]. Specifically, given a sphere with a radius  $r$ , we

employ the widely adopted dart-throwing method [12] to sample colors within it until no more discernible colors can be sampled. Here, a discernible color means that the perceptual differences between it and those sampled colors exceed a threshold of 10, which is consistent with the threshold we used in Sec. 4.2.1. Our experimental results indicate a roughly linear relationship between the maximal number of possible colors  $n$  and the square of the radius:  $n \propto r^2$ . Therefore, we introduce constraints that  $r_i/r_j = \sqrt{n_i}/\sqrt{n_j}$ , where  $r_i$  and  $r_j$  denote the radii of two spheres associated with two parent classes, and  $n_i$  and  $n_j$  denote the number of child classes within those spheres. Finally, the radii of the spheres are determined as the maximal radii that adhere to both  $d_{12} - r_1 - r_2 > \max(r_1, r_2)$  and  $r_1/r_2 = \sqrt{n_1}/\sqrt{n_2}$ .

## 5 EVALUATION

### 5.1 Quantitative Evaluation

**Datasets.** We evaluated the quality of different color assignment results across 12 datasets, which have been widely used in recent visualization research [64, 77]. Six of them (MNIST [28], Animals [14], Indian Food [42], Isolet [18], Texture [1], Clothing [66]) are flat datasets with a moderate number of classes, ranging from 10 to 26. These datasets are used to evaluate different flat color assignment methods. The remaining six datasets (Food101 [3], Flowers102 [44], Stanford Cars [26], Caltech256 [21], NABirds [58], ImageNet1k [16]) have a larger number of classes, ranging from 101 to 1000. These datasets are used to evaluate different hierarchical color assignment methods. For datasets with a pre-existing class hierarchy, such as ImageNet, we directly used their hierarchies. For the other datasets, we applied the commonly used hierarchical k-means method [65] to build the class hierarchy in a top-down manner. More details of these datasets are given in the supplemental material.

We used these datasets to create three types of visualization: scatter-plots (point-based) [8, 64], parallel coordinates (line-based) [19], and grid visualizations (area-based) [9, 77], which cover the primary types of visualization techniques for data analysis. We also evaluated the generated color assignment results on their own, *i.e.*, considering the color palettes solely without integrating them into specific visualizations. When evaluating the flat color assignment methods, we assigned colors to all classes. When evaluating hierarchical color assignment methods, we simulated how users explore the hierarchy by randomly choosing a subtree for expansion.

**Baseline methods and our method variations.** We chose four state-of-the-art color assignment methods for comparison. Palettaylor [36] and Color Crafting [53] are two representative flat color assignment methods. Palettaylor focuses on optimizing color discrimination and incorporates spatial distribution to improve this aspect further, while Color Crafting ensures basic color discrimination and focuses more on color harmony. Tree Colors [55] and Cuttlefish [60] are two representative hierarchical color assignment methods. Tree Colors is a static method that assigns colors to all classes, while Cuttlefish is a dynamic method that only assigns colors to visible classes during exploration. Since Palettaylor and Color Crafting are flat color assignment methods, we extended them to support hierarchical color assignment. Palettaylor is capable of generating color assignment results within a specified hue range. Therefore, we employed our dynamic color selection method to select the hue ranges for child classes and then used Palettaylor to generate color assignment results within each range. Color Crafting generates a sequence of colors with a similar hue but different lightness

Table 2: Comparison of our method with the representative color assignment methods, with the best in **bold** and the second best underlined. The values in *gray* indicate that the colors cannot be easily distinguished.

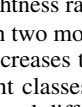
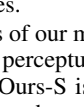
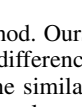
PD: perceptual difference, ND: name difference, CL: chroma-lightness, BHDI: balanced harmony-discrimination index, SS: silhouette score, DR: distance ratio.

Methods	Flat color assignment					Hierarchical color assignment						
	Discriminability		Harmony		BHDI	Discriminability		Harmony		BHDI	Alignment with hierarchy	
	PD	ND	Hue	CL		PD	ND	Hue	CL		SS	DR
Palettailor	19.419	0.913	0.296	0.377	4.441	8.202	0.508	0.838	0.579	3.253	0.582	0.883
Color Crafting	2.549	0.261	<b>1.000</b>	<b>1.000</b>	2.777	4.176	0.243	<b>0.997</b>	0.998	2.898	0.642	0.930
Tree Colors	6.168	0.848	0.608	<b>1.000</b>	3.909	0.226	0.066	0.978	<b>1.000</b>	2.133	<b>0.785</b>	<b>0.994</b>
Cuttlefish	9.944	0.813	0.702	<b>1.000</b>	4.323	4.615	0.382	0.606	<b>1.000</b>	2.832	0.715	0.907
Ours-D	<b>23.194</b>	<b>0.921</b>	0.876	0.955	<u>5.992</u>	<b>16.579</b>	<b>0.736</b>	0.984	0.810	4.926	<u>0.740</u>	0.945
Ours-S	<u>23.070</u>	<u>0.920</u>	<u>0.893</u>	<u>0.962</u>	<b>6.002</b>	<u>16.482</u>	<u>0.699</u>	<u>0.985</u>	0.927	<b>4.958</b>	<u>0.740</u>	<u>0.946</u>

levels. We first chose colors with different hues for the classes at the top level. Subsequently, we estimated the lightness range for child classes and used Color Crafting to generate color assignment results for each child class within these specified lightness ranges.

We compared these methods with two modes of our method. Ours-D is the difference mode, which increases the perceptual difference between colors of spatially adjacent classes. Ours-S is the similarity mode, which reduces the perceptual difference between colors of similar classes.

**Evaluation criteria.** We used seven measures to evaluate the quality of color assignment results from three perspectives: discriminability, harmony, and alignment with hierarchical structures. Discriminability is evaluated using perceptual difference (PD) and name difference (ND), which have been introduced in Sec. 4.2.1. Harmony is measured using hue harmony (Hue) and chroma-lightness harmony (CL), which have been introduced in Sec. 4.2.2. While discriminability and harmony are both crucial, they often conflict with each other. For example, using identical colors maximizes harmony, but minimizes discriminability. To address this, we design a new measure, the balanced harmony-discrimination index (BHDI), to combine these four measures and give a more comprehensive one. Specifically, we follow the hyperparameters we set in Sec. 4 to combine them, *i.e.*,  $0.1 \times PD + 2.0 \times ND + \text{Hue} + \text{CL}$ . Alignment with hierarchical structures is evaluated using silhouette score (SS) and distance ratio (DR), which are not used in our optimization process. SS [50] measures the compactness and separability of colors among child classes. A higher value indicates that child classes of the same parent class have more similar colors, and those of different parent classes have more distinct colors. DR quantifies color similarity between child classes and their parent classes. It calculates the color distance from a child class to the closest class at the parent level and compares it to the distance to its actual parent class. A ratio closer to 1 indicates a better alignment between child and parent classes. A more thorough comparison with additional measures is available in the supplemental material.

**Results.** Table 2 presents the comparison results between the baseline methods and our methods. From the flat color assignment results, it can be seen that our methods perform best in terms of discriminability. In particular, our methods achieve a perceptual difference of 23, which significantly exceeds the threshold of 10 for high accuracy in color discrimination [5]. In terms of harmony, our methods rank second in hue harmony, with only Color Crafting ahead. However, it should be noted that Color Crafting’s high hue harmony comes at the sacrifice of discriminability, which is evidenced by its worst discriminability (2.549 in perceptual difference and 0.261 in name difference) and the visualization results (*e.g.*, colors  in Fig. 7(b)). Our methods fall behind in terms of chroma-lightness harmony compared to Color Crafting, Tree Colors, and Cuttlefish. The main reason is that these methods apply either a fixed template or a strict constraint to determine the chroma and lightness of colors. This results in worse discriminability, such as the colors  in Fig. 7(c) and the colors  in Fig. 7(d). Instead, our methods sacrifice a little chroma-lightness harmony to allow greater variation to enhance discriminability. Despite this, our methods perform best in terms of BHDI and produce visually appealing results (Figs. 7(e) and (f)). These demonstrate the strength of our methods in

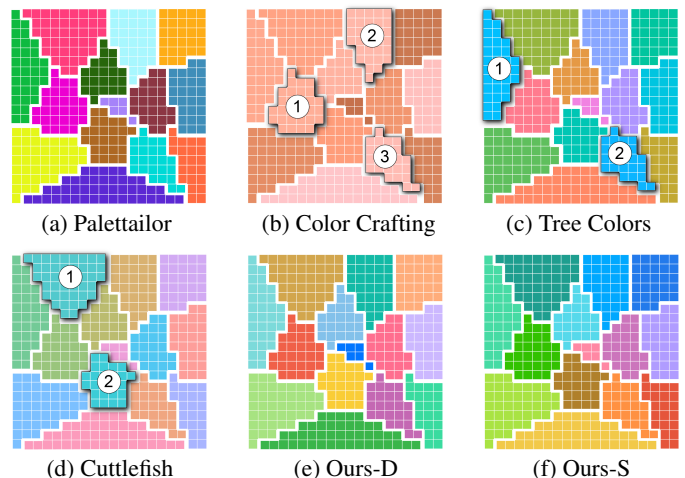
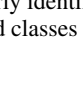
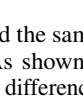


Fig. 7: Flat color assignment results generated by different methods.

balancing discriminability and harmony, which are both indispensable in visualization. This is further confirmed by the user study results in Sec. 5.2.1 as our methods are highly favored by the experts.

For hierarchical color assignment, the results on discriminability, harmony, and BHDI are similar to those for flat color assignment. Therefore, we focus on how well these methods align with hierarchical structures. As shown in Table 2, our methods rank second among all methods, slightly behind Tree Colors. The primary reason for Tree Colors’ high performance is that it enforces the colors of child classes within a small range around the colors of parent classes, which achieves a tight alignment with hierarchical structures. However, this results in extremely poor discriminability between colors, which is even worse than Color Crafting (0.226 *vs.* 4.176 in perceptual difference and 0.066 *vs.* 0.243 in name difference). The visualization results in Fig. 8E also show that Tree Colors generates almost identical colors for five child classes of the same parent (). In contrast, our methods generate colors that are both discriminable and clearly identifiable as belonging to the same parent class for these five child classes ( in Fig. 8A).

## 5.2 User Study

We also conducted a user study to capture human preferences for different color assignment methods.

### 5.2.1 User Study Design

**Methods.** In our user study, we used the same four baseline methods as in our qualitative evaluation. As shown in Table 2, only minor differences are observed between the difference mode (Ours-D) and the similarity mode (Ours-S). This prompts us to choose one to simplify the comparison process for experts. We chose the similarity mode because it shows slightly better performance in BHDI and alignment with hierarchical structures, and this mode is more suitable for data analysis in practice [67].

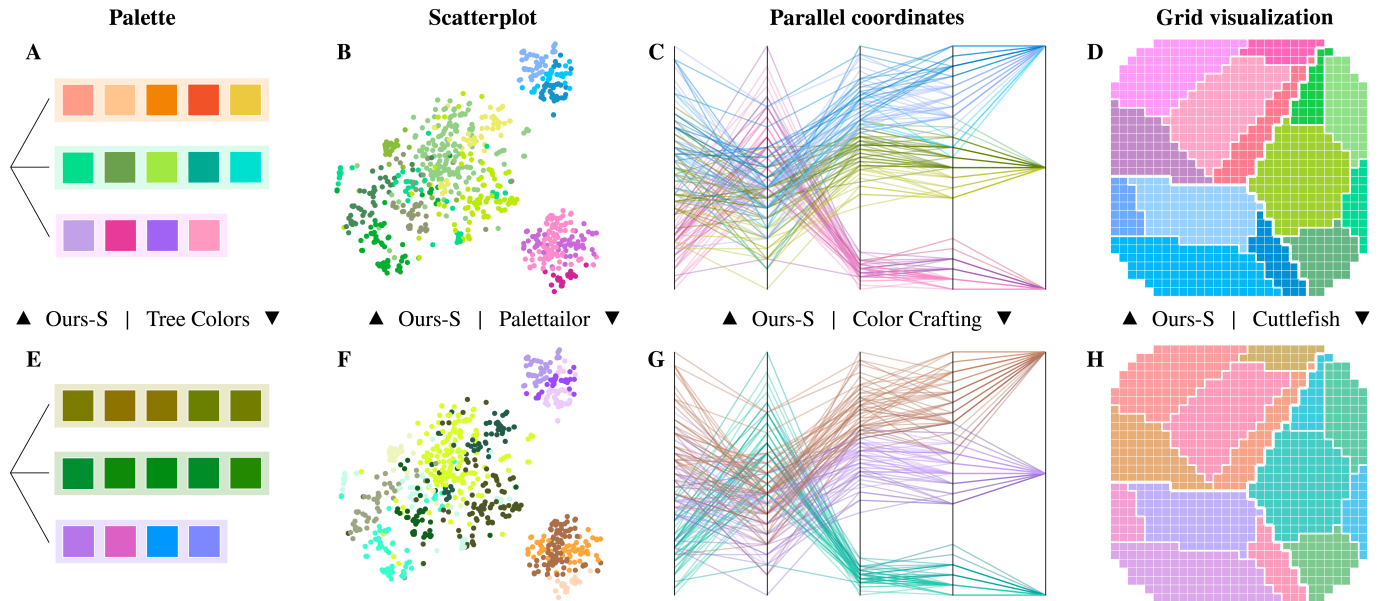


Fig. 8: Hierarchical color assignment results generated by different methods and displayed in different visualization types.

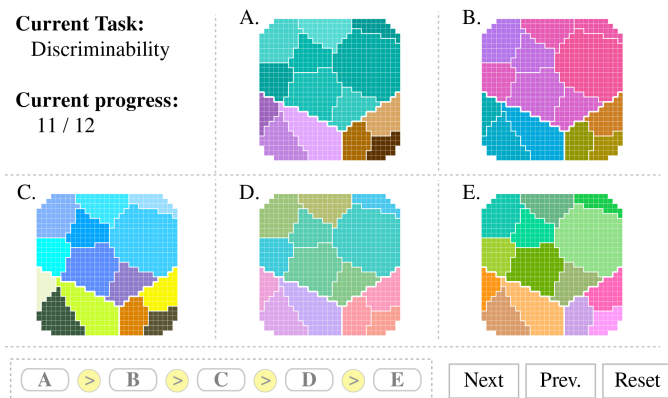


Fig. 9: The interface of the user study.

**Experts.** We recruited 20 experts for our study, including 12 males and 8 females. Nine of them specialize in color design, and the remaining eleven are experts in information visualization. All of them confirmed their expertise in color design and data visualization. None of them reported any color deficiency. Upon completion, each expert was rewarded with a \$30 gift card.

**Study procedure.** The user study consists of three tasks that require the experts to rank the color assignment results based on discriminability, harmony, and alignment with hierarchical structures, respectively. Before each task, we provided a brief overview of these concepts. Following this, experts were required to complete 12 trials for each task. In each trial, experts ranked the color assignment results on a web-based interface (Fig. 9). They could take a brief break after completing each task. Upon completing all trials, they were asked to complete a questionnaire, which collected their personal information and detailed feedback on how they evaluated the five color assignment results depicted in Fig. 9. Each study lasted 40-60 minutes.

**Conditions and design.** Similar to the quantitative evaluation, we included palettes, scatterplots, parallel coordinates, and grid visualizations in our user study to cover different types of visualization. This diverse inclusion ensures that our evaluation results are more robust and reliable. However, to avoid overburdening experts, we did not include all the 12 datasets used in the quantitative evaluation, as doing so would result in a total of 144 trials (3 tasks  $\times$  4 visualizations  $\times$  12 datasets). This would require an excessive amount of time from our experts. Instead, we considered three representative scenarios in

data exploration: 1) examining a higher hierarchical level with a balanced subclass distribution (high-bal), which is the most common case in data exploration; 2) examining a higher hierarchical level with an imbalanced subclass distribution (high-imb), which sometimes happens during exploration; 3) examining a lower hierarchical level (low), where the available color range will become much narrower. This makes our findings more convincing and applicable to a wide range of data exploration scenarios. Consequently, each expert went through 36 trials (3 tasks  $\times$  4 visualizations  $\times$  3 scenarios).

### 5.2.2 Result Analysis

**Overall comparison.** First, we computed the average rank for each color assignment method. If an expert ranked two methods equally, their ranks were set as their average rank. For example, if the ranking is  $A=B>C=D=E$ , the ranks of A and B will be  $(1+2)/2=1.5$ , and the ranks of C, D, and E will be  $(3+4+5)/3=4$ . Next, we conducted Friedman tests and pairwise Wilcoxon signed-rank tests to compare the ranks of different methods. The statistical test results and the box plots are presented in Fig. 10. The Friedman test results indicate significant differences among the methods in discriminability ( $\chi^2(4) = 42.68, p < 0.0001$ ), harmony ( $\chi^2(4) = 41.96, p < 0.0001$ ), and alignment with hierarchical structures ( $\chi^2(4) = 45.70, p < 0.0001$ ). Therefore, we focus on pairwise comparisons in subsequent analysis. Regarding discriminability, our method achieves an average rank of 1.64, which is significantly better than all the baseline methods. Regarding harmony, our method achieves a comparable performance with Cuttlefish (an average rank of 2.04 vs. 2.19), and they are both significantly better than the other three methods. Regarding the alignment with hierarchical structures, our method ranks highest on average (1.70), which is significantly better than the other methods. These findings are consistent with the ones in the quantitative evaluation that our method is effective in balancing discriminability and harmony, while also maintaining better alignment with hierarchical structures. Further details of the distribution of the ranks are summarized in the supplemental material.

**Detailed comparison on different scenarios.** We further investigate whether the ranking results vary across different exploration scenarios. As shown in Fig. 10, the average ranks of different methods do not significantly change between the scenario of exploring a higher hierarchical level with a balanced subclass distribution and that with an imbalanced class distribution (high-bal vs. high-imb). However, in the exploration of lower hierarchical levels, both Color Crafting and Tree Colors exhibit a notable drop in discriminability. This is because, at lower levels, they produce color assignment results within a narrower color range. Moreover, the experts ranked these two methods



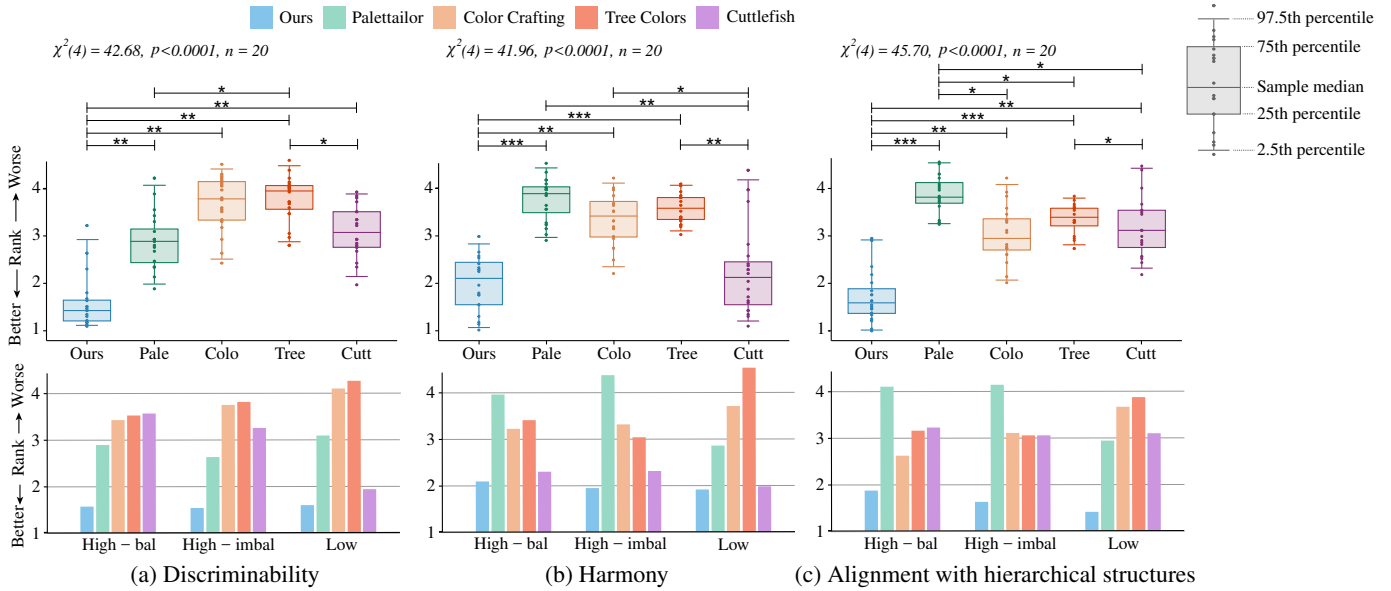


Fig. 10: User study results in three tasks. Top: Friedman tests and pairwise Wilcoxon signed-rank tests on five methods. Bottom: Detailed comparison of their rankings under three exploration scenarios. \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ , and \*\*\* indicates  $p < 0.001$ .

lower in harmony. This is because they believed harmony is built on contrast, colors that are similar but not identical are perceived as more harmonic. In contrast, both our method and Cuttlefish maintain good performance in discriminability and harmony, which is achieved by dynamically expanding the color range in lower levels. Notably, our method achieves better alignment with hierarchical structure compared to Cuttlefish because of the improved dynamic color range selection method. This improvement ensures a closer correspondence between the colors assigned to child classes and their respective parent classes, thereby improving overall hierarchical coherence.

## 6 EXPERT FEEDBACK AND DISCUSSION

After the user study, we summarized expert feedback on their evaluation of different color assignment methods in terms of discriminability, harmony, and alignment with hierarchical structures. To gain deeper insights, we conducted semi-structured interviews with eight experts. In each interview, we first presented their ranking results and encouraged them to explain their choices, especially those where their opinions differed from the majority. After going through all the results, we held an open discussion to collect more feedback on our color assignment method and explore potential opportunities for enhancement. The duration of each discussion was around 50-60 minutes.

**Discriminability.** When discussing how they compare different color assignment results in terms of discriminability, many experts highlighted the important role of hue and saturation. This led to a higher ranking for options C/D/E compared to A/B in Fig. 9. Two experts also pointed out that to develop a color assignment method for a wide audience, it is necessary to consider users with color vision deficiency (CVD). Given the flexibility of our method, we are optimistic about its potential to generate CVD-friendly color assignment results. The key lies in selecting an appropriate color range and refining the calculation of color discriminability. Some research on CVD models [4, 38] can be integrated to achieve this goal. In addition, Stone *et al.* [54] explored the impact of mark size on discriminability. A promising research direction involves incorporating the findings into the optimization process to tailor color assignment results for various visualization types, including scatterplots, line charts, and grid visualizations.

**Harmony.** Many experts pointed out that they would first rank color assignment results with extremely high saturation or luminance, such as Fig. 8F and Fig. 9C, as the least harmonic. They also noted that harmony decreased when colors were too similar to be easily distinguished [48]. This explains the lower ranking of Color Crafting and Tree Colors in lower hierarchical levels. Although our method and

Cuttlefish scored lower in the harmony metric, they were still preferred by experts. Moreover, several experts pointed out that personal preferences and cultural differences also affect the perception of color harmony [45, 49, 73]. For example, one expert favored warm colors with low saturation, such as the results generated by Cuttlefish (Fig. 8H). Our methods can integrate these individual and cultural preferences by modifying the color ranges or including customized harmonic patterns. However, it is difficult for end users to accurately describe their preferences in terms of color ranges. It remains an opportunity to explore how to collect user feedback (like/dislike) and model user preferences.

**Consistency with hierarchical structures.** The experts noted that when comparing different methods, they focused more on the perceptual difference between the colors of parent classes and child classes. A smaller difference indicates a clearer parent-child relationship. They also highlighted the critical role of hue difference between different parent classes, which would significantly facilitate the accurate identification of parent-child relationships. This justifies our additional emphasis on the hue channel in the dynamic color range selection process. Moreover, some experts prioritized discriminability between the colors of parent classes and would tolerate relatively smaller discriminability between the colors of child classes of the same parent. Other experts held the opposite view and believed that maintaining discriminability between the colors of child classes was more important. Therefore, it is necessary to provide a customizable trade-off between these two goals. Currently, our method ensures that the gap between two color ranges for the child classes of different parents should be greater than the radii of each color range. Users can modify this requirement to achieve a smaller or larger gap to adapt to their needs.

## 7 CONCLUSION

We develop a dynamic color assignment method that simultaneously considers discriminability, harmony, and spatial distribution. It also dynamically assigns colors based on user exploration and aligns them with hierarchical structures. Our method starts by generating a discriminable and harmonic color assignment result for top-level classes within the full color range. When users zoom in on a region for detailed analysis, our method selects an appropriate color range for the child classes based on the colors of the selected parent classes. Subsequently, our method generates the color assignment result for child classes within the selected color ranges and ensures discriminability and harmony. The effectiveness of our method is demonstrated through a quantitative evaluation and a user study, which highlights its capability in generating high-quality dynamic color assignment results.

## SUPPLEMENTAL MATERIALS

All supplemental materials are available on OSF at [https://osf.io/e4b5u/?view\\_only=68cc67c194c443b498bd2545ef551faa](https://osf.io/e4b5u/?view_only=68cc67c194c443b498bd2545ef551faa), released under a CC BY 4.0 license. In particular, they include (1) dataset information, (2) the running time of our color assignment methods, (3) additional color assignment results, (4) additional quantitative evaluation results, (5) additional user study results, and (6) the video.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under grants U21A20469, 61936002, grants from the Institute Guo Qiang, THUIBCS, and BLBCI. The authors would like to thank Jun Yuan, Zhen Li, and Duan Li for their valuable contributions to the discussions, and Yiwei Hou for her assistance in voicing our video.

## REFERENCES

- [1] D. Arthur, Asuncion. UCI machine learning repository. <https://sci2s.ugr.es/keel/dataset.php?cod=72>, 2007. Last accessed 2024-8-1. 6
- [2] Y. Bengio, J. Louradour, R. Collobert, and J. Weston. Curriculum learning. In *Proceedings of the International Conference on Machine Learning*, pp. 41–48. ACM, Montreal, 2009. doi: 10.1145/1553374.1553380 2, 5
- [3] L. Bossard, M. Guillaumin, and L. Van Gool. Food-101—mining discriminative components with random forests. In *Proceedings of the European Conference on Computer Vision*, pp. 446–461. Springer International Publishing, Cham, 2014. doi: 10.1007/978-3-319-10599-4\_29 6
- [4] H. Brettel, F. Viénot, and J. D. Mollon. Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America A*, 14(10):2647–2655, 1997. doi: 10.1364/JOSAA.14.002647 9
- [5] A. Brychtová and A. Çöltekin. The effect of spatial distance on the discriminability of colors in maps. *Cartography and Geographic Information Science*, 44(3):229–245, 2017. doi: 10.1080/15230406.2016.1140074 3, 7
- [6] C. Chen, J. Chen, W. Yang, H. Wang, J. Knittel, X. Zhao, S. Koch, T. Ertl, and S. Liu. Enhancing single-frame supervision for better temporal action localization. *IEEE Transactions on Visualization and Computer Graphics*, 30(6):2903–2915, June 2024. doi: 10.1109/tvcg.2024.3388521 3
- [7] C. Chen, Y. Guo, F. Tian, S. Liu, W. Yang, Z. Wang, J. Wu, H. Su, H. Pfister, and S. Liu. A unified interactive model evaluation for classification, object detection, and instance segmentation in computer vision. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):76–86, 2024. doi: 10.1109/TVCG.2023.3326588 5
- [8] C. Chen, Z. Wang, J. Wu, X. Wang, L.-Z. Guo, Y.-F. Li, and S. Liu. Interactive graph construction for graph-based semi-supervised learning. *IEEE Transactions on Visualization and Computer Graphics*, 27(9):3701–3716, 2021. doi: 10.1109/TVCG.2021.3084694 6
- [9] C. Chen, J. Yuan, Y. Lu, Y. Liu, H. Su, S. Yuan, and S. Liu. OoDAnalyzer: Interactive analysis of out-of-distribution samples. *IEEE Transactions on Visualization and Computer Graphics*, 27(7):3335–3349, 2021. doi: 10.1109/TVCG.2020.2973258 6
- [10] H. Chen, W. Chen, H. Mei, Z. Liu, K. Zhou, W. Chen, W. Gu, and K.-L. Ma. Visual abstraction and exploration of multi-class scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1683–1692, 2014. doi: 10.1109/TVCG.2014.2346594 2
- [11] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y.-Q. Xu. Color harmonization. In *ACM SIGGRAPH 2006 Papers*, pp. 624–630. ACM, New York, 2006. doi: 10.1145/1179352.1141933 2, 3
- [12] R. L. Cook. Stochastic sampling in computer graphics. *ACM Transactions on Graphics*, 5(1):51–72, 1986. doi: 10.1145/7529.8927 5, 6
- [13] K. R. Coombes, G. Brock, Z. B. Abrams, and L. V. Abruzzo. Polychrome: Creating and assessing qualitative palettes with many colors. *bioRxiv*, 1(1):303883, 2018. doi: 10.1101/303883 5
- [14] A. Corrado. Animals-10 dataset. <https://www.kaggle.com/datasets/alessiocorrado99/animals10/>, 2019. Last accessed 2024-8-1. 6
- [15] J. Delon, A. Desolneux, J. L. Lisani, and A. B. Petro. Automatic color palette. In *Proceedings of the IEEE International Conference on Image Processing*, vol. 2, pp. II–706. IEEE, Genoa, 2005. doi: 10.1109/ICIP.2005.1530153 2
- [16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 248–255. IEEE, Miami, 2009. doi: 10.1109/CVPR.2009.5206848 6
- [17] H. Fang, S. Walton, E. Delahaye, J. Harris, D. Storchak, and M. Chen. Categorical colormap optimization with visualization case studies. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):871–880, 2017. doi: 10.1109/TVCG.2016.2599214 1, 2
- [18] M. Fandy and R. Cole. Spoken letter recognition. In *Proceedings of Advances in Neural Information Processing Systems*, pp. 220–226. Morgan Kaufmann, Denver, 1990. 6
- [19] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner. Hierarchical parallel coordinates for exploration of large datasets. In *Proceedings of IEEE Visualization*, pp. 43–508. IEEE, San Francisco, 1999. doi: 10.1109/VISUAL.1999.809866 1, 2, 6
- [20] C. C. Gramazio, D. H. Laidlaw, and K. B. Schloss. Colorgorical: Creating discriminable and preferable color palettes for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):521–530, 2017. doi: 10.1109/TVCG.2016.2598918 1, 2, 3, 5
- [21] G. Griffin, A. Holub, and P. Perona. *Caltech-256 Object Category Dataset*. California Institute of Technology, 2007. 6
- [22] C. G. Healey. Choosing effective colours for data visualization. In *Proceedings of IEEE Visualization*, pp. 263–270. IEEE, San Francisco, 1996. doi: 10.1109/VISUAL.1996.568118 2
- [23] J. Heer and M. Stone. Color naming models for color selection, image editing and palette design. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1007–1016. ACM, Austin, 2012. doi: 10.1145/2207676.2208547 2, 4
- [24] Y. Jiang, P. Xu, C. Zhang, H. Fu, H. Lau, and W. Wang. Region-aware color smudging. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–13, 2024. doi: 10.1109/TVCG.2024.3374210 3
- [25] N. Kita and K. Miyata. Aesthetic rating and color suggestion for color palettes. *Computer Graphics Forum*, 35(7):127–136, 2016. doi: 10.1111/CGF.13010 2
- [26] J. Krause, M. Stark, J. Deng, and L. Fei-Fei. 3d object representations for fine-grained categorization. In *Proceedings of the IEEE International Conference on Computer Vision workshops*, pp. 554–561. IEEE, Sydney, 2013. doi: 10.1109/ICCVW.2013.77 6
- [27] C. Lara-Alvarez and T. Reyes. A geometric approach to harmonic color palette design. *Color Research & Application*, 44(1):106–114, 2019. 4, 5
- [28] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791 6
- [29] S. Lee, M. Sips, and H.-P. Seidel. Perceptually driven visibility optimization for categorical data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(10):1746–1757, 2013. doi: 10.1109/TVCG.2012.315 2, 3, 5
- [30] K. Li, J. Li, Y. Sun, C. Li, and C. Wang. Color assignment optimization for categorical data visualization with adjacent blocks. *Journal of Visualization*, 26(4):917–936, 2023. doi: 10.1007/s12650-022-00905-z 3
- [31] J. Lin, P. Xiao, Y. Fu, Y. Shi, H. Wang, S. Guo, Y. He, and T.-Y. Lee. C3 assignment: Camera cubemap color assignment for creative interior design. *IEEE Transactions on Visualization and Computer Graphics*, 28(8):2895–2908, 2022. doi: 10.1109/TVCG.2020.3041728 2
- [32] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu. Towards better analysis of deep convolutional neural networks. *IEEE transactions on visualization and computer graphics*, 23(1):91–100, 2017. 5
- [33] S. Liu, Y. Liang, and A. Gitter. Loss-balanced task weighting to reduce negative transfer in multi-task learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 9977–9978. AAAI Press, Honolulu, 2019. doi: 10.1609/AAAI.V33I01.33019977 5
- [34] S. Liu, M. Tao, Y. Huang, C. Wang, and C. Li. Image-driven harmonious color palette generation for diverse information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):1–16, 2022. doi: 10.1109/TVCG.2022.3226218 1, 2, 5
- [35] S. Liu, W. Yang, J. Wang, and J. Yuan. *Visualization for Artificial Intelligence*. Springer, 2024. 1
- [36] K. Lu, M. Feng, X. Chen, M. Sedlmair, O. Deussen, D. Lischinski, Z. Cheng, and Y. Wang. Palettaior: Discriminable colorization for categorical data. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):475–484, 2021. doi: 10.1109/TVCG.2020.3030406 1, 2, 3, 4, 5, 6
- [37] K. Lu, K. Reda, O. Deussen, and Y. Wang. Interactive context-preserving color highlighting for multiclass scatterplots. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–15. ACM, Hamburg, 2023. doi: 10.1145/3544548.3580734 2

- [38] G. M. Machado, M. M. Oliveira, and L. A. Fernandes. A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1291–1298, 2009. doi: 10.1109/TVCG.2009.5344284
- [39] D. Mahapatra and V. Rajan. Multi-task learning with user preferences: Gradient descent with controlled ascent in pareto optimization. In *Proceedings of the International Conference on Machine Learning*, pp. 6597–6607. PMLR, Virtual Event, 2020. doi: 10.1137/2020m1
- [40] Y. Matsuda. Color design. *Asakura Shoten*, 2(4):10, 1995. doi: 10.1109/TVCG.2009.5344284
- [41] B. A. Maxwell. Visualizing geographic classifications using color. *The Cartographic Journal*, 37(2):93–99, 2000. doi: 10.1179/0008704.37.2.p93
- [42] Mehta, Anshul. Indian food. <https://www.kaggle.com/datasets/anshulmehtakaggl/themassiveindianfooddataset>, 2020. Last accessed 2024-8-1. doi: 10.1109/TVCG.2020.9344284
- [43] P. Nardini, M. Chen, M. Böttinger, G. Scheuermann, and R. Bujack. Automatic improvement of continuous colormaps in euclidean colorspace. *Computer Graphics Forum*, 40(3):361–373, 2021. doi: 10.1111/cgf.14313
- [44] M.-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *Proceedings of the Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, pp. 722–729. IEEE, Bhubaneswar, 2008. doi: 10.1109/ICVGIP.2008.476
- [45] Z. O’Connor. Colour harmony revisited. *Color Research & Application*, 35(4):267–273, 2010. doi: 10.1002/col.20624
- [46] L.-C. Ou, P. Chong, M. R. Luo, and C. Minchew. Additivity of colour harmony. *Color Research & Application*, 36(5):355–372, 2011. doi: 10.1002/col.20624
- [47] L.-C. Ou and M. R. Luo. A colour harmony model for two-colour combinations. *Color Research & Application*, 31(3):191–204, 2006. doi: 10.1002/col.20208
- [48] L.-C. Ou, Y. Yuan, T. Sato, W.-Y. Lee, F. Szabó, S. Sueeprasarn, and R. Huertas. Universal models of colour emotion and colour harmony. *Color Research & Application*, 43(5):736–748, 2018. doi: 10.1002/col.22243
- [49] S. E. Palmer and K. B. Schloss. An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences*, 107(19):8877–8882, 2010. doi: 10.1073/pnas.0906172107
- [50] P. J. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20(1):53–65, 1987. doi: 10.1016/0377-0427(87)90125-7
- [51] V. Setlur and M. C. Stone. A linguistic approach to categorical color assignment for data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):698–707, 2016. doi: 10.1109/TVCG.2015.2467471
- [52] G. Sharma, W. Wu, and E. N. Dalal. The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, 2005. doi: 10.1002/col.20070
- [53] S. Smart, K. Wu, and D. A. Szafrir. Color crafting: Automating the construction of designer quality color ramps. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1215–1225, 2020. doi: 10.1109/TVCG.2019.2934284
- [54] M. Stone, D. A. Szafrir, and V. Setlur. An engineering model for color difference as a function of size. In *Proceedings of the Color and Imaging Conference*, pp. 253–258. Society for Imaging Science and Technology, Boston, 2014. doi: 10.2352/CIC.2014.22.1.ART00045
- [55] M. Tennekes and E. de Jonge. Tree colors: Color schemes for tree-structured data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2072–2081, 2014. doi: 10.1109/tvcg.2014.2346277
- [56] C. Tseng, G. J. Quadri, Z. Wang, and D. A. Szafrir. Measuring categorical perception in color-coded scatterplots. In *proceedings of the 2023 CHI conference on human factors in computing systems*, pp. 1–14. ACM, Hamburg, 2023. doi: 10.1145/3544548.3581416
- [57] E. R. Tufte. Envisioning information. *Optometry and Vision Science*, 68(4):322–324, 1991. doi: 10.1016/1045-926X(92)90022-E
- [58] G. Van Horn, S. Branson, R. Farrell, S. Haber, J. Barry, P. Ipeirotis, P. Perona, and S. Belongie. Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 595–604. Boston, 2015. doi: 10.1109/CVPR.2015.7298658
- [59] N. Waldin, M. Le Muzic, M. Waldner, E. Gröller, D. Goodsell, A. Ludovic, and I. Viola. Chameleon: Dynamic color mapping for multi-scale structural biology models. In *Eurographics Workshop on Visual Computing for Biomedicine*, vol. 2016. Eurographics Association, Bergen, 2016. doi: 10.2312/vcbm.20161266
- [60] N. Waldin, M. Waldner, M. Le Muzic, E. Gröller, D. S. Goodsell, L. Autin, A. J. Olson, and I. Viola. Cuttlefish: Color mapping for dynamic multi-scale visualizations. *Computer Graphics Forum*, 38(6):150–164, 2019. doi: 10.1111/cgf.13611
- [61] L. Wang, J. Giesen, K. T. McDonnell, P. Zolliker, and K. Mueller. Color design for illustrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1739–1754, 2008. doi: 10.1109/TVCG.2008.118
- [62] Y. Wang, X. Chen, T. Ge, C. Bao, M. Sedlmair, C.-W. Fu, O. Deussen, and B. Chen. Optimizing color assignment for perception of class separability in multiclass scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):820–829, 2019. doi: 10.1109/TVCG.2018.2864912
- [63] M. Wu, Y. Sun, and S. Jiang. Adaptive color transfer from images to terrain visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):1–16, 2023. doi: 10.1109/TVCG.2023.3295122
- [64] S. Xiang, X. Ye, J. Xia, J. Wu, Y. Chen, and S. Liu. Interactive correction of mislabeled training data. In *2019 IEEE Conference on Visual Analytics Science and Technology*, pp. 57–68. IEEE, Vancouver, 2019. doi: 10.1109/VAST47406.2019.8986943
- [65] J. Xiao, J. Lu, and X. Li. Davies bouldin index based hierarchical initialization k-means. *Intelligent Data Analysis*, 21(6):1327–1338, 2017. doi: 10.1109/ICCVPR.2015.7298885
- [66] T. Xiao, T. Xia, Y. Yang, C. Huang, and X. Wang. Learning from massive noisy labeled data for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2691–2699. IEEE, Boston, 2015. doi: 10.1109/CVPR.2015.7298885
- [67] Y. Xue, P. Paetzold, R. Kehlbeck, B. Chen, K. C. Kwan, Y. Wang, and O. Deussen. Reducing ambiguities in line-based density plots by image-space colorization. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):825–835, 2024. doi: 10.1109/TVCG.2023.3327149
- [68] W. Yang, Y. Guo, J. Wu, Z. Wang, L.-Z. Guo, Y.-F. Li, and S. Liu. Interactive reweighting for mitigating label quality issues. *IEEE Transactions on Visualization and Computer Graphics*, 30(3):1837–1852, 2024. doi: 10.1109/TVCG.2023.3345340
- [69] W. Yang, Z. Li, M. Liu, Y. Lu, K. Cao, R. Maciejewski, and S. Liu. Diagnosing concept drift with visual analytics. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pp. 12–23, 2020. doi: 10.1109/vast50239.2020.00007
- [70] W. Yang, M. Liu, Z. Wang, and S. Liu. Foundation models meet visualizations: Challenges and opportunities. *Computational Visual Media*, 10(3):399–424, May 2024. doi: 10.1007/s41095-023-0393-x
- [71] W. Yang, X. Wang, J. Lu, W. Dou, and S. Liu. Interactive steering of hierarchical clustering. *IEEE Transactions on Visualization and Computer Graphics*, 27(10):3953–3967, 2021. doi: 10.1109/TVCG.2020.2995100
- [72] W. Yang, X. Ye, X. Zhang, L. Xiao, J. Xia, Z. Wang, J. Zhu, H. Pfister, and S. Liu. Diagnosing ensemble few-shot classifiers. *IEEE Transactions on Visualization and Computer Graphics*, 28(9):3292–3306, 2022. doi: 10.1109/TVCG.2022.3182488
- [73] K. Yokosawa, N. Yano, K. B. Schloss, L. R. Prado-León, and S. E. Palmer. Cross-cultural studies of color preferences: Us, japan, and mexico. *Journal of Vision*, 10(7):408–408, 2010. doi: 10.1167/10.7.408
- [74] J. Yuan, S. Xiang, J. Xia, L. Yu, and S. Liu. Evaluation of sampling methods for scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1720–1730, 2021. doi: 10.1109/TVCG.2020.3030432
- [75] L.-P. Yuan, Z. Zhou, J. Zhao, Y. Guo, F. Du, and H. Qu. Infocolorizer: Interactive recommendation of color palettes for infographics. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4252–4266, 2022. doi: 10.1109/TVCG.2021.3085327
- [76] Q. Zheng, M. Lu, S. Wu, R. Hu, J. Lanir, and H. Huang. Image-guided color mapping for categorical data visualization. *Computational Visual Media*, 8(4):613–629, 2022. doi: 10.1007/s41095-021-0258-0
- [77] Y. Zhou, W. Yang, J. Chen, C. Chen, Z. Shen, X. Luo, L. Yu, and S. Liu. Cluster-aware grid layout. *IEEE Transactions on Visualization and Computer Graphics*, 30(1), 2024. to be published. doi: 10.1109/TVCG.2023.3326934