Categorical colormap optimization with visualization case studies

This item was submitted to Loughborough University’s Institutional Repository by the/an author.

Citation: FANG, H. ... et al., 2016. Categorical colormap optimization with visualization case studies. IEEE Transactions on Visualization and Computer Graphics, 23(1), pp. 871 - 880.

Additional Information:

- Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Metadata Record: https://dspace.lboro.ac.uk/2134/36576

Version: Accepted for publication

Publisher: © Institute of Electrical and Electronics Engineers (IEEE)

Please cite the published version.
1 INTRODUCTION

Color is one of the most prevalent visual channels used in visualization because of its merits in visual search and easy implementation [49]. Categorical colormaps (or qualitative colormaps) are commonly featured in visualization for mapping data values to distinct colors, allowing users to identify categories of values quickly through association [4]. The choice of colors can affect the users’ ability to learn and remember the mapping [3, 35], and more importantly, their ability to differentiate the colors perceptually when such a mapping is used in a visualization [2, 22, 9].

Colormap design has been a subject of much study. Many schemes for choosing colors exist, including historical convention [6], system recommendations [1, 47, 20], and metaphoric colors [31]. There have been a number of design guidelines, e.g., [21, 7, 43, 39]. This paper is concerned with a classic question in specifying a categorical colormap, that is, when a visualization designer has selected a set of colors based on one of the above schemes or by following certain guidelines, how can this colormap be further optimized?

Lee et al. [29] provided an optimization method based on the visual saliency of each point against its surroundings in a visualization. While this is a useful solution for optimizing a specific visualization, it is not suitable for a system to generate numerous visualizations for different datasets using a consistent colormap. We thus devised a new algorithmic approach that enables users to define several types of constraints, for instance, (i) to fix a certain number of colors (e.g., for background colors), and (ii) to limit the variation of certain properties of the original colors (e.g., to maintain the intended color semantics).

In terms of measuring color differences, the merits of perceptually-uniform color spaces (e.g., CIE L∗u∗v∗ and L*a*b*) over commonly-used color spaces (e.g., RGB, HSB) is well documented in the color literature. There are several proposed metrics for measuring color distance in a perceptually-uniform space, including CIE76, CIE94, CIEDE2000 and CMC84 [22]. There is a general consensus that CIEDE2000 outperforms other metrics (e.g., [51]). We thus decided to use CIEDE2000 (cf. CIE76 in [29]), and focus our effort on the algorithmic issues in optimization.

In this work, we examine three algorithms in terms of convergence, speed and deployment. In particular, we provide a technical solution to address the phenomena of local maxima in optimization because the distance space based on CIEDE2000 is not totally convex. For example, as shown in Fig. 7(a), the lines of the London underground were introduced gradually into a colormap. Because of the need for maintaining the historic color mappings for passengers, this has created challenges for making new lines differentiable. While the Nelder-Mead simplex method has a low computational cost, optimization is often impeded by local maxima. Meanwhile the technique of genetic algorithm can address this issue, but it requires user-defined constraints to prevent the loss of the original semantic association.

We also report a case study where seismologists needed to specify the categories of depths of seismic events, and would like to improve an existing colormap for visualizing the depths of seismic events in two visual representations. As both visual representations are typically used by the data analysts hundreds of times a day, it is important for the colors in the colormap to be easily memorable and distinguishable. This work has enabled a scientific approach to the design of colormaps in this application.

Our contributions include (i) a novel approach for optimizing a categorical colormap algorithmically while enabling users to define initial metaphoric colors and set constraints for the optimization (Sections 3 and 5); (ii) a study of three optimization algorithms in this context and compared their relative merits in terms of convergence, speed, and deployment (Sections 4); (iii) two case studies that demonstrate the practical usability of our approach (Section 7); and (iv) a web-based application to enable wider access to some of the presented techniques.

2 RELATED WORK

In this section, we first give a brief survey of color designs in visualization. We then give an overview of perceptually-uniform color spaces, and optimization in such spaces.

Color Designs in Visualization. Color is one of the most powerful visual channels in visualization, as it facilitates pre-attentive visual search and pop-out effect. In this respect, there is a common consensus that among the four commonly-used visual channels, color precedes size, shape and orientation. This has been well-supported in the psychology literature (e.g., [50, 38, 44, 37]). Bertin [4] considered that color is associative, selective and ordered. Ware [48] conducted empirical studies on colormaps in the context of visualization. Borgo et al. [5] and Demiralp et al. [8] studied the distance between visual objects featuring integrated visual channels (including colors) using crowdsourced experiments.
Visualization textbooks (e.g., by Stone [43], Ware [49], Munzner [33], and Tufte [45]) offer guidance about colormaps. Other sources of guidance include [12, 21, 7, 6, 39]. Several tools have been developed for helping users design colormaps. Bergman et al. [1] developed a rule-based approach taking spatial frequency into account. Meier et al. [32] developed an interactive tool for users to “mix and organize colors; explore color combinations; and solicit historical, theoretical, or expert sources.” Harrower and Brewer [20] developed the popular web-based tool, ColorBrewer, providing pre-defined colormaps for sequential, diverging and qualitative forms of mapping.

Colors have a profound role in history, culture and everyday life [13, 19, 15]. Lin et al. [31] used automated web searches to determine the association between a given set of terms and the colors in representative images. Kuhn et al. [27] proposed an automated method for recoloring images for dichromats.

**Perceptually-uniform Color Spaces.** For many decades, much effort has been made to develop *perceptually uniform color spaces* where every pair of equal distances measured corresponds with a pair of equally-perceived color differences. The development of Lab color spaces (lightness and two color dimensions) saw the evolution from Hunter’s Lab space in 1948, to CIE 1976 L*a*b* color space, to the CIE 1994 and CIEDE2000 distance functions [23, 26, 22]. Today the difference measure $\Delta E$ in CIEDE2000 remains as the industrial standard. Although CIEDE2000 in general offers a better distance measure than other formulae, it is important to note that it does not guarantee perceptual uniformity in all conditions (e.g., varying light sources or materials). There are several parameters that should be tuned for each individual condition. The CIEDE2000 standard provides parameters for only two conditions, none of which represent computer displays. We will address this issue in Section 4.

Several works in the field of visualization considered the need for optimizing a pre-defined colormap. Kindlmann et al. [25] proposed a method for colormap optimization by using faces as measuring templates. Wang et al. [47] describe a rule-based system for optimizing the colors chosen for 2D and 3D visualizations, accommodating factors such as color discrimination, translucency and labeling. Chuang et al. [11] proposed to optimize colormaps for lowering energy consumption. Other optimization works, which do not involve a perceptually uniform space, include [40, 30].

Lee et al. [29] presented a color optimization algorithm that maximizes the perceptual distances among a set of categorical colors. Their work is based on CIE L*a*b* space, for which the authors pointed out its limitation, suggested a remedy of using CIE94 and CIEDE2000, and reasoned about the algorithmic difficulty in using CIEDE2000. Our work differs from [29] in several ways: (i) We focused on the optimization of a color palette rather than individual visualization; (ii) We enabled users to define constraints in a more meaningful way and demonstrated the use of such constraints through a real-world application; (iii) We made use of CIEDE2000, studied a number of optimization algorithms and identified their relative merits in addressing the algorithmic difficulty; and (iv) We provided a web application to make this technique available to a wider range of users.

3  PROBLEM STATEMENT

Let $C = \{c_1, c_2, \ldots, c_n\}$ be a set of colors. Let $D(c_i, c_j)$ be a distance function for measuring the difference between two colors. We assume that $D$ satisfies all essential conditions for a metric; i.e., non-negativity, identity of indiscernibles, symmetry, and triangle inequality. We denote the minimal distance among all pairs of colors as $D_{\min}(C)$, i.e.,

$$D_{\min}(C) = \min_{i,j} D(c_i, c_j), \quad \forall i, j (i \neq j)$$

(1)

The general optimization problem is to find a set of color transformations, $\Phi_i, i = 1, 2, \ldots, n$, such that the set of transformed colors $C' = \{\Phi_1(c_1), \Phi_2(c_2), \ldots, \Phi_n(c_n)\}$ has reached maximal value of $D_{\min}(C')$.

The problem, however, is not as simple as the above general formulation. There are following additional considerations:

- **Color Space** — There are several color spaces, such as RGB, HSL, HSV, and CMYK, which are widely understood by users and supported by software for drawing, presentation, and image manipulation. Other color spaces, such as CIE XYZ, CIE L*a*b*, CIE UVW, and CIE L*a*b*′ were developed according to measurements of human color perception. There are also color spaces designed for television broadcasts and other industrial applications. The distance calculation can take place in any of such spaces, but will yield different results.

- **Distance Metrics** — Given colors are defined using 3- or 4-tuples, many distance functions can be candidates for such a function. These include the commonly-used Euclidean distance, the Manhattan distance, and their generalization, the Minkowski distance. There are also many other candidates, such as Chebyshev distance and normed vector space. Some color spaces, such as CIE L*a*b*, have predefined distance functions, while others allow for more flexible interpretations.

- **Constraints** — In practice, not all colors need to be or should be transformed. For example, the background color of a display canvas, and/or the foreground color for labeling may be fixed. In many cases, the users may wish to restrict the amount of changes allowed during color transformation. Such constraints normally relate to the semantics of a colormap, and may be defined for one or a few specific colors, or for all colors.

In both our case studies, to be detailed in Section 7, there is a combination of requirements for the above considerations. On one hand, the distance calculations must take place in a perceptually-uniform color space, while on the other hand, the constraints must be defined in a color space familiar to users. On one hand, optimization algorithms prefer to work on a convex distance space, while on the other hand the property of perceptually-uniformity may require a distance function that is not convex. On one hand, a color scheme may be used for displaying foreground information, and on the other hand, a metaphorically similar scheme is required for displaying background information. All these make the color optimization a non-trivial undertaking.

4  GENERAL OPTIMIZATION

Global optimization techniques have been used previously in the field of visualization (e.g., [36, 29]). Fig. 1 shows four iteration steps in optimizing the distances among four colors. The four colors are shown in a special palette where each color (circle) is overlaid on all other colors (strips). The pairwise fitness diagram (e) shows the CIEDE2000 distance of each pair of colors over time. The multi-band colormap is defined by minimal, average, and maximal distances ($D_{\min}, D_{\avg}, D_{\max}$) at $t_0$, together with absolute 0-distance and the black-white distance ($D_{BW}$). The four ranges are:

- $[0, D_{\min}]$ in Wine – At $t_0$, it shows the minimum. After $t_0$, any such color is highly undesirable and is expected to disappear.
- $[D_{\min}, D_{\avg}]$ in Yellow – At $t_0$, it shows any distance equal to or below the average. Colors in this range indicate undesirable distances, and should ideally be reduced over time.
- $[D_{\avg}, D_{\max}]$ in Green – At $t_0$, it shows any distance above the average at $t_0$. Colors in this range indicate desirable distances, and more of such distances are expected to appear after $t_0$.
- $[D_{\max}, D_{BW}]$ in Grey – As colors in this range are above the maximum at $t_0$, they appear only after $t_0$. The more, the better.

In our work, we focus on optimizing a pre-defined initial categorical colormap such that the input colormap is improved to a level that cannot be further improved algorithmically. We consider the following three goals that potentially can be in conflict: (i) the perceptual distance between the colors should be maximized according to a distance metric; (ii) the user should be able to define constraints to limit undesired changes; and (iii) the fixed background color and foreground colors (if any) should be factored into the optimization process.
4.1 Perceptual Color Metrics

There are many metrics for measuring color difference. Perception-based metrics attempt to simulate the color differences perceived by the human visual system. Most of these metrics perform in the CIE \( L^*a^*b^* \) color space. CIE1976 is the simplest metric measuring Euclidean distances in \( L^*a^*b^* \). More advanced metrics, CIE1994 and CIEDE2000, were introduced to address perceptual non-uniformity in \( L^*a^*b^* \). Among CIE standards, CIEDE2000 [10, 42] is currently considered to be the closest to human perception. Its formula is widely available in the literature.

CIEDE2000 has three pre-determined parameters, \( K_L, K_C \) and \( K_H \), to balance the weights of the color components for different visual media, with the default value 1 for each of them. One common adjustment is to set \( K_L = 2 \) for textiles, and \( K_L = 1 \) for graphic arts [26]. However, computer displays are neither textiles nor graphic arts. Different sets of \( K_L, K_C \) and \( K_H \) for computer displays were proposed in the literature, e.g., [52, 51, 53]. After comparing their experimental conditions and results, we set \( K_L = 0.725 \), \( K_C = 1 \) and \( K_H = 1 \) in our calculations.

In general, the larger the distance between a pair of colors, the more perceptually distinguishable they are. This notion is however probabilistic. This is partly because the perception of color differences varies from person to person and from chart to chart [12]. It is also because the CIEDE2000 distance space is highly nonlinear from the perspective of the RGB space. The transformation from RGB to CIE \( L^*a^*b^* \) introduces the first distortion [18], and the distance calculation introduces the second distortion since CIEDE2000 is by definition not a Euclidean distance metric [10].

The visible color space is bounded. The more colors a colormap tries to accommodate, the smaller \( D_{min} \) will be. Green-Armytage proposed 26 possibly distinguishable colors after a series of empirical studies. The minimal pairwise distance is 6.86 (CIEDE2000 is unit-less). Among 325 pairs of distances, 3.69% below 20, and 6.77% below 25. Since several colors among the 26 colors are very similar, it suggests that [20, 25] might be the borderline zone.

4.2 Fitness Function

Our optimization strategy is to increase the minimal perceptual distance among different pairs of colors in the colormap. Although some other works, e.g., [29], used a strategy to maximize the average distance, we believe that improving the color pairs with the shortest distances yields greatest benefit to visualization. Thus a solution with a higher minimum and a lower average is more preferable than that with a lower minimum and a higher average. Based on Eq. 1, given \( m \) intermediate states \( C_1, C_2, \ldots, C_m \), the fitness function selects \( k < m \) states such that \( D_{min}(C_{selected}) > D_{min}(C_{unselected}) \). At the final iteration, the state with the maximal \( D_{min} \) is selected as the solution.

5 Constrained Optimization

When optimizing a colormap algorithmically, users may potentially find themselves with a colormap that is perceptually optimized, but suboptimal for its application. For example, many categorical colormaps are designed purposely to associate colors with symbolic or metaphoric meanings. When new colors are introduced to an existing colormap, it is usually desirable to maintain the consistency and memorability of the old colors. The two case studies in Section 7 exemplify these two scenarios.
Comparing with Fig. 2, we can see that the shifting towards purple was
placed only in terms of saturation and luminance in the HSL space.
One may wish to fix the hue of the colors to be optimized in order to
control of the search range by recalculating the actual range based on the
current state of the solution. As a result, the adaptive range provides a
mechanism to transform colors during optimization at a gentle pace.
An example of using adaptive range is shown in Fig. 5. When using
the technique of genetic algorithm (GA) without the adaptive range control,
the green in the initial palette in Fig. 5 (a) becomes a blue color in the resultant
palette in (b), while the blue becomes green. With an adaptive range constraint,
the problem of swapping in GA can be alleviated. Fig. 5(c) shows the result
by setting $\Delta$ to 10% for each of the H, S, and V color components.
The undesired swapping of green and blue is eradicated.

6 Optimization Algorithms

Given the problem statement in Section 3, a full search for an optimal
solution is of a complexity $O((k_1 k_2 k_3)^n)$, where $n$ is the number of
colors, and $k_1, k_2$ and $k_3$ are the numbers of valid values of the
two color components. For example, for a typical display in the RGB
space, we have $k_1 = k_2 = k_3 = 256$, and $256^n$ combinations in total.
The basic hill climbing algorithm avoids a full search by following the
most promising path from an initial color. There are $m = 3n$ variables
for all the components of the $n$ colors. For each iteration, we may consider three possible options for transforming a color component:
no change ($\Phi(x) = x$), increment ($\Phi(x) = x + \delta$), or decrement
($\Phi(x) = x - \delta$), where $\delta > 0$. Hence, each iteration would need to
evaluate $3^n = 3^{3n}$ optional transformations for $n$ colors. Therefore
the full search and hill climbing are both intractable. To reduce the

color symbolism and metaphors, and logological association), this
type of constraint can also enable users to have more control over
the solution to be generated to suit a particular, artistic style, display
medium, or display environment. For example, overly saturated col-
ors are often avoided by visualization designers typically for aesthetic
reasons. In order to avoid saturated colors at $t_{100}$ in Figs. 1, 2 and
3, one may set a desired saturation range $[0.1, 0.4]$. Under this fixed
range constraint, the optimization process first transforms the initial
value of the saturation component into the range using Eq. 2, and then
starts the iteration as the normal constrained optimization. The results
are shown in Fig. 4.

Adaptive Range. The adaptive range is defined by $\Delta$, the percentage
of changes allowed at each iteration. Let $x_i$ be the value of an arbitrary
color component $X$ at the $n$th iteration. The transformed color $x'_i$
is constrained by the range $[x_i(1 - \Delta), x_i(1 + \Delta)]$, together with $[X_{\min}, X_{\max}]$;
the overall range for $X$. After each iteration, the algorithm takes con-
trol of the search range by recalculating the actual range based on the
current state of the solution. As a result, the adaptive range provides a
mechanism to transform colors during optimization at a gentle pace.

Fig. 5. This example shows the comparison of optimization with and
without adopting the adaptive range. Without using the adaptive range control
when running the GA algorithm, the result in (b) has a color swap
effect while the adaptive range control makes smooth evolution.

Fig. 4. Optimization with a range constraint on saturation, which is set
between $[0.1, 0.4]$. The initial palette in (a) shows the results after the
normalization of the initial colors using Eq. 2. The optimization is also
constrained by a white background.
With the input color as the initial best solution and the first option to explore. The initial temperature $T_0$ is set to 100 (i.e., 100% probability). At the $i^{th}$ iteration, $T_i = T_0 \cdot 0.95^i$. We made use of MATLAB’s `simulannealbnd`, which also enables the control of the exploration boundary for constrained optimization.

**Genetic Algorithm (GA)** has been widely used in many applications [17]. The optimization process is defined as an evolution of a collection of $\omega$ chromosomes, each containing an optional transformation. In each generation (i.e., iteration), the $\omega$ chromosomes are evaluated using the fitness function $D_{\text{min}}$, $\nu \ll \omega$ best chromosomes are selected to progress to the next generation. Meanwhile, $\omega - \nu$ pairs of distinct parents are probabilistically chosen from the $\omega$ chromosomes. The chromosomes with better fitness values have more chances to be chosen. These parents result in $\omega - \nu$ children to progress to the next generation. The transformation contained in each child chromosome is a linear combination of the transformations in its two parent chromosomes. We used the `ga` function in MATLAB, which also includes a mutation operation for introducing the additional diversity in evolution. After some experimentation, we set $\omega = 80$, $\nu = 2$, and $\text{mutation} = 0.03$. The initial 80 chromosomes contain the set of $n$ input colors, and 79 sets obtained by uniformly sampling the HSL space, within the fixed or adaptive range if set by the user.

**Comparison of the Algorithms.** We compare the three algorithms in terms of the minimal distance $D_{\text{min}}$ reached, and the speed for achieving convergence. Using a range of tests with different settings of colors and constraints, the comparison was fairly conclusive. Fig. 6 shows the testing results of the two examples in Figs. 3 and 4.

Fig. 6 shows a phenomenon of GA $> ((\text{NM} \geq \text{SA}) \cup (\text{NM} \leq \text{SA}))$ in terms of $D_{\text{min}}$. This was quite typical in our tests. In general, GA tended to achieve the largest $D_{\text{min}}$, while the ordering between NM and SA was less consistent.

In order to compare the computational costs in a consistent manner, we test the three algorithms by making sure that they will evaluate the same number of options in each iteration. In NM and SA, we grouped every 80 consecutive iterations into a macro iteration in order to be consistent with GA ($\omega = 80$). As exemplified by Figs. 6(d, h), NM converged much more quickly than GA and SA, while GA was slightly faster than SA in most cases.

Furthermore, NM and SA were more likely to be trapped in a local maximum. GA alleviated this problem, but would often deviate from the input colors, such as changing the order of color symbolism in Fig. 5(b). The adaptive range constraint introduced in Section 5 can prevent GA from such deviations.

7 Application Case Studies

This work was initially motivated by the need to design a number of categorical colormaps in a visualization system for visualizing seismological data. For this scientific visualization application, it was difficult to resolve conflicts of color symbolism by simply selecting each colormap using software such as ColorBrewer. To test the utility of the technique developed in a more general scenario, we applied it to the classic problem of visualizing the London Tube Map.

7.1 Case Study #1: The London Tube Map

The London tube map is iconic in the history of visualization. In 1863, the world saw its first underground railway when the Metropolitan Railway opened in London. Since then it has been developed into an integrated underground and overground network with 14 lines. When these lines were gradually introduced, as shown in Fig. 7(a), the colormap of the London Tube Map had to assign distinct colors to new lines while maintaining the old color symbolism. There is no doubt that the designers of every evolution of the tube map made good efforts to optimize the colormap, perhaps based mainly on their artistic intuition. Nevertheless, some colors in the current standard colormap for computer screens [16] have rather short distances in terms of CIEDE2000. For example, as shown in Fig. 7(a), the shortest CIEDE2000 distance is 15.6 between DLR and Waterloo & City.

We may wish to optimize the colormap with a need for preserving the historical color symbolism. We can, for instance, fix the colors of

![Fig. 6. Comparing the three algorithms in terms of the convergence of $D_{\text{min}}$ and their computational cost. In (a-d), the setting of Fig. 3 is used, and in (e-h), the setting of Fig. 4 is used. $D_{\text{min}}$ is the minimal distance among all color pairs. A larger value indicates a better solution.](image-url)
Fig. 7. The London underground has massively expanded over one and a half centuries. (a) Many pairwise CIEDE2000 distances in the TfL (Transport for London) standard screen colormap [16] are rather close. (b) To maintain color symbolism for all lines, we fix the colors of pre-1980 lines, and optimize those of post-1980 lines using the Nelder-Mead simplex method. (c) We can further increase the distances by relaxing the constraints for all lines. (d) Using the genetic algorithm method, the colors of post-1980 lines can break out of local maximum regions. (e) Slightly relaxing the colors of all lines allows us to address the short distances among pre-1980 colors. (f) Meanwhile, the TfL standard web colormap also features some short CIEDE2000 distances. Seven shortest distances in each map are labelled. Colors being displayed may be affected by device-dependent settings.

Table 1. Metaphoric colors defined before color optimization.

<table>
<thead>
<tr>
<th>Depth (km)</th>
<th>Metaphor (Alliterative Association)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shallow Ranges</td>
<td>California (Cherry, Crimson); Shallowness (Scarlet)</td>
</tr>
<tr>
<td>15, 35</td>
<td>Mexico (Mango)</td>
</tr>
<tr>
<td>35, 70</td>
<td>Bismarck Sea (Blue)</td>
</tr>
<tr>
<td>Intermediate Ranges</td>
<td>Hindu (Honey)</td>
</tr>
<tr>
<td>70, 160</td>
<td>Andean (Amethyst)</td>
</tr>
<tr>
<td>160, 250</td>
<td>Japanese (Jasmine); Deep (Daisy, Daffodil)</td>
</tr>
<tr>
<td>500, 700</td>
<td>Fiji (Forest Green)</td>
</tr>
<tr>
<td>≥ 700</td>
<td>Possibly Problematic (Pink)</td>
</tr>
</tbody>
</table>

7.2 Case study #2: Visualizing Seismological Data

The International Seismological Centre (ISC) has been collecting and verifying earthquake data for more than a century. As part of a visual system for supporting the workflow of a team of data analysts, there are some seven views, each requiring a different colormap. As the data analysts use the system throughout each working day, they need to be able to remember color mappings in each view with little need for checking the legends. Color symbolism thus plays a critical role in supporting familiarization and memorization. Here we describe how color optimization helps the design of two related views, where depths are categorized into different bands.

Hypocentres Overview. This view, which is also referred to as the Historical Seismicity Map, enables data analysts to examine whether the depth estimations reported by different agencies for a seismic event are consistent with the historical events in the region. Fig. 8(a) shows a view of Hypocentres Overview, with the old colormap that was originally designed by a former staff member many years ago. While the use of a categorical colormap is essential for rapid comparison, there is no record about its design rationale. For scientific and operational reasons, domain experts decided to modify the categorization scheme (from 9 to 8 bands), and took the opportunity to define a new colormap featuring more meaningful color symbolism. Note that the partitioning of the numerical range was purposely designed to be non-uniform, reflecting both geological and precisional factors.

Table 1 shows a set of modified depth bands, and initial metaphoric colors chosen by the domain experts and visualization designers in a brainstorming meeting. Statistics about different regions at each depth band were made available during the meeting. Each color metaphor was based on the name of a geographical region, which is well-known for having seismic events with depth in a specific band. The name of a region is then mapped to a color using alliterative associations (a form of logology). In some cases, additional semantic meanings, such as “shallow” and “deep”, provide additional alliterative associations. For example, historically California had many earthquakes at a shallow depth. Hence the shallowest band (< 15 km) is exemplified by this region, and is mapped to red through alliterative associations cherry/imson/california and scarlet/shallow. Because earthquakes in a shallow range usually cause more damages, the red color also draws attention to such events. Fig. 8(b) shows the same view with the metaphorical colormap resulting from the brainstorming meeting.

In addition, there are three background colors, RGB(255,255,255),
Fig. 8. Four versions of a historical seismicity map (of a region near Tonga) exemplify the evolution from the original 9-band map in (a), to an 8-band map with metaphoric colors in (b), to a map optimized without hue constraint in (c), and finally to a map in (d) optimized with hue constraint to preserve color metaphors. By comparing (b) and (d), we can observe that the band [250, 500) becomes more distinguishable from the white background and the band [15, 35) becomes more separable from [0, 15) and [70, 160). Meanwhile, the loss of some color symbolism in (c), e.g., for bands [0, 15) and [15, 35), is avoided in (d).

RGB(240,240,240), and RGB(0,0,0), depicting ocean, land, and nation borders respectively. As the colormap is expected to be used for any arbitrary area on the earth, it was difficult to optimize the 11 colors intuitively by inspecting numerous views of different areas. Figs. 8(c, d) show the same view with two colormaps resulting from GA optimization. In the case of (c), we did not set any constraints except fixing the three background colors. This allows colors to be transformed more freely. We can see that the hues of some colors may have shifted too much to maintain the original alliterative associations. For example, the first band becomes more of an orange, connecting to neither california or shallow. We therefore added a ±5% constraint on each hue component of the eight metaphoric colors. As shown in Fig. 8(d),
Fig. 9. The depthmap view of two seismic events: an event at an intermediate depth (left column), and an event at a large depth (right column). (a) The traditional design could not facilitate a rapid perception of different y-scales between the two visualizations. (b) The background color-coded using the colormap in Fig. 8(d) solved the problem in (a), but the colors were not optimal. (c) A visually more pleasing colormap was obtained using the optimization technique with user-defined constraints.

the new colormap maximized the color separation while maintaining the original metaphors and alliterative associations.

Hypocentre Depth View. This view displays a specially designed bar chart for showing the estimated depths of an event reported by different agencies. It is designed for observing potential inconsistencies among different reports. As shown in Fig. 9, the reported depths were listed in an ascending order from left to right. There are two types of depth values, fixed or free. The former, denoting a heuristic solution, is depicted by a black circle at the bottom of a depth bar. The latter, denoting a computational solution, is depicted by error bars showing the standard deviation.

Domain experts had a critical requirement that the y-axis should be scaled according to the maximum values reported, and horizontal grid lines should correspond to the eight depth bands as discussed above. While the scaling might appear to be counterintuitive to visualization designers (cf. [46, 34]), there were strong scientific reasons for this requirement in this application. For example, estimations for an event at a shallow depth are normally more accurate than those for a deep event. It is necessary to see the smaller depth values in detail, while a coarse visual approximation of larger depth values is acceptable. The same numerical difference between two depth values may suggest high inconsistency in shallow ranges, but fairly acceptable in deep ranges. The same standard deviation would also be judged differently in different ranges. In addition, it was desirable to help data analysts relate a depth value to its categorization (i.e., one of the eight bands).

The initial sketch of the depth view, as illustrated by two examples in Fig. 9(a), used dashed lines as the separator of depth bands. We were unconvinced about its effectiveness as it was difficult to relate a depth value to a depth category. We then color-coded each band of the background canvas using the colormap in Fig. 8(d). The results are shown in Fig. 9(b). Domain experts found this approach highly effective for viewing a depth value with its spatial mapping of length as well as its categorization. However, the dark background colors were thought to be “very hard” on eyes, especially considering the very frequent use of this view in daily tasks. Using GA, we transformed the colormap in Fig. 8(d) to a pastel version by constraining hues with ±5% variation, saturations in the range of [0.3, 0.5], and luminance in the range of [0.8, 0.9]. The results, as shown in Fig. 9(c), achieved better contrast between the background colors and the foreground black for depth bars. In addition, the color-coded background offers another important visual cue for data analysts to judge the level of scaling of the y-axis. Comparing Fig. 9(c), a user can unconsciously become
The merits of CIE L*a*b* and CIEDE2000 have been extensively studied and widely accepted in color science. Hence our evaluation focused on specific sets of colors generated using the aforementioned optimization techniques. As this work was carried out at the ISC where one visualization researcher was embedded in a team of data analysts, we adopted user-centered design and agile development for all views in the visualization system. There were weekly meetings involving the lead analyst, the database manager, and visualization researchers. Major landmarks in design and development were examined in detail through brainstorm meetings or software testing involving other data analysts.

For the colormaps used in the aforementioned two views, we collect feedback from the data analysts through questionnaires and face-to-face meetings. The data analysts confirmed the following merits of color optimization:

- Using metaphors provided significant benefits, including being easy to remember, and easy to articulate when training new analysts and communicating with the external stakeholders.
- When the team assigned the metaphoric meaning to the categorical colormap, there were doubts among the users about the quality of the manually-defined colormap (Fig. 8(b)). After applying the color optimization, the users felt more confident about the quality of the new colormap (Fig. 8(d)).
- Among the three design options shown in Fig. 9), the users overwhelmingly preferred the version in the third row. They welcomed the colormaps featuring the same metaphor but different saturation and luminance, enabling them to relate the Hypocentres Overview and Hypocentre Depth View with ease.
- The optimization tool was also used in designing the colormaps for other views in the system.

8 Colourmap Hospital - A Web-based Tool

In order to make the developed technique available to a wider community, we have produced a web-based tool called Colourmap Hospital. The application enables a user to specify a set of colors with constraints, optimize these colors, and view the optimization results. Fig. 10 shows three screenshots of this tool, which is available at [http://bit.ly/cmhosp](http://bit.ly/cmhosp).

Users specify the colors to be optimised in either RGB (numeric or hexadecimal) or HSL, and can switch between spaces at any time. A live swatch view showing pairwise comparisons of the colors is updated on the right-hand side during specification. Constraints can be specified with two mechanisms. (i) An individual constraint can be specified by clicking the lock icon to the left of a color. Through three sliders, a user can set +/− ranges around the current hue, saturation and luminance values. (ii) The globally-applicable constraint, once enabled, applies to all colors that do not have an individual constraint. A user can set this constraint by entering a percentage value, and assigning it to hue, saturation and/or luminance. For example, entering 10% and checking the “hue” checkbox places a 36-degree hue boundary around all colors without any individual constraint.

The usability of the tool was evaluated through a consultancy meeting with 10 participants, all of whom had experience of using colormaps in visualization. Following a brief introduction, participants performed five optimization tasks using the tool. A short survey and discussions confirmed the usability strongly in terms of distinguishable colors, and moderately in terms of pleasing or satisfactory results and easy to use. Participants also suggested design aspects to be improved. Further details are given in the supplementary material.

The optimization itself is computed in a remote Python server application. Because of the need to provide color optimization at an interactive rate through a web interface, the tool currently supports only the Nelder-Mead simplex method, since its computational cost is significantly lower than the other two algorithms tested. The optimization is computed in CIE L*a*b* with CIEDE2000. Testing results have confirmed that the Python implementation is consistent with the MATLAB implementation described in Section 6.

9 Conclusions

Designing a categorical colormap is an elementary yet important task in data visualization. There is no doubt that colormaps designed manually by experienced visualization researchers and visual artists or based the recommendation of software such as ColorBrewer are adequate for many applications. Nevertheless, designers can benefit from an optimization tool in a number of scenarios. For example, the tool can be used to refine a design sketch computationally, enabling the designers to focus more effort on other requirements for a colormap (e.g., preserving historical or creating new color symbolisms). It can also be used to check and update the existing colormaps used in various systems. Our two application case studies demonstrate the utility of such a tool.

It is necessary to note that a numerically optimized colormap is not necessarily the most suitable colormap. As shown in Fig. 9, the colormaps in (b) represents an optimized colormap with better distance measure than those in (c). Hence, the humans’ judgement about the quality in different contexts is still essential in practice. Nevertheless, with the tool, the colormaps in (b) can be transformed to those in (c) much more easily than with manual transformation.

We wish to further develop the techniques of color optimization. In particular, there are needs for optimizing continuous colormaps.