

面向数据可视化的色彩计算

曾琼¹⁾, 汪云海¹⁾, 屠长河¹⁾, 陈宝权^{1,2)}

¹⁾(山东大学计算机科学与技术学院 青岛 266000)

²⁾(北京大学前沿计算研究中心 北京 100871)

(qiong.zn@sdu.edu.cn)

摘要: 色彩在数据可视化中发挥着重要作用, 是人们读取数值、感知趋势、发现异常的关键视觉编码元素。然而, 色彩设计过程需要较多人为干预, 存在低效耗时、难以拓展且易产生色彩的误用、滥用等问题。近年来兴起的数据可视化评估理论为解决上述问题提供了思路, 即将色彩设计表达为在数据、用户、任务、设备等可视化应用环境下的最优决策。文中回顾了数据可视化领域色彩设计相关工作, 主要包括色彩设计准则及其量化表达、自动色彩设计(色彩选择、色彩优化、色彩生成、色彩迁移)以及自适应色彩设计, 论述了现有色彩计算研究对于探索智能化色彩设计的意义与启发, 总结并分析了该领域未来发展趋势。

关键词: 数据可视化; 色彩设计; 色彩计算

中图法分类号: TP391.41 DOI: 10.3724/SP.J.1089.2020.18487

Color Computing in Data Visualization

Zeng Qiong¹⁾, Wang Yunhai¹⁾, Tu Changhe¹⁾, and Chen Baoquan^{1,2)}

¹⁾(School of Computer Science and Technology, Shandong University, Qingdao 266000)

²⁾(Center on Frontiers of Computing Studies, Peking University, Beijing 100871)

Abstract: Color plays a significant role in data visualization, which is one of key visual channels for quantity estimation, pattern perception and anomaly detection. However, color design is often a trial-and-error process, which is inefficient, hardly to extend and easily causes misleading effects. Recently, the development in visualization evaluation inspires the automated color design. The intelligent color design can be considered as the optimal decision under the conditions of specific data, users, tasks and even devices. In this paper, we will review the references of color design in data visualization, including color design guidelines and perceptual models, automatic color design (colormap selection, colormap optimization, colormap generation and color retargeting), and adaptive color design in dynamic environment. We will discuss the role of existing color design research in motivating intelligent color design algorithms, conclude and analyze the future development in color computing.

Key words: data visualization; color design; color computing

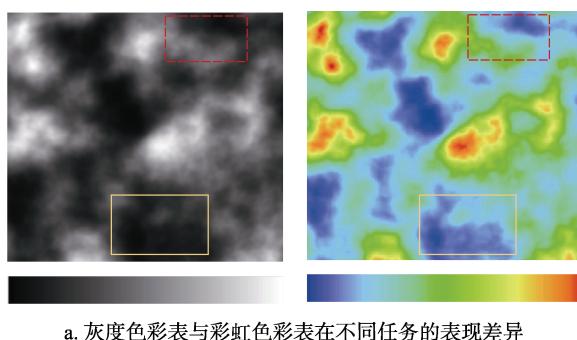
色彩学是一门集光学物理学、美学、心理学等
为一体的交叉学科, 从三色理论(trichromacy theory)

到色彩数值度量(color measurement), 色彩空间
(color space)到色彩外观感知(color appearance), 都

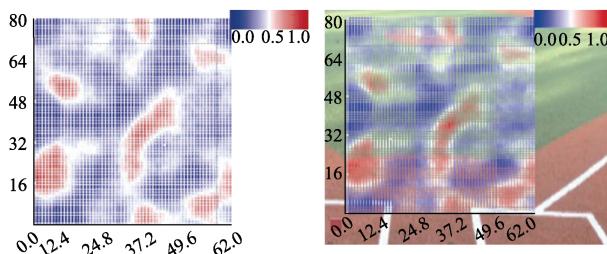
收稿日期: 2020-07-09; 修回日期: 2020-07-22. 基金项目: 国家自然科学基金(61861136012, 61772315). 曾琼(1987—), 女, 博士, 副研究员, CCF 会员, 主要研究方向为数据可视化; 汪云海(1984—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为数据可视化与可视分析; 屠长河(1968—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为计算机图形学与计算几何; 陈宝权(1969—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为计算机图形学与数据可视化。

具有深厚的理论积累。色彩在可视化中的应用可追溯到 18 世纪，主要用于统计学图形化表达。如英国的弗罗伦斯·南丁格尔曾基于色彩以及圆形图绘制了著名的“东部军队死亡原因统计图”，分析出克罗米亚战争的死亡原因主要为战场外感染疾病以及护理不当。与光学物理学、美学等领域关注于色彩自身属性(如波段、美感等)不同，数据可视化中的色彩设计以让用户高效地理解数据、发现规律、探索任务为目标，侧重于挖掘色彩与数据、任务、设备等可视化应用环境之间的关联关系。

然而，在实际可视化应用中，色彩设计的过程通常为迭代试错的过程(trial-and-error process)，即设计师或用户迭代尝试不同色彩设计并基于个人主观感知判断其合理性。由于可视化中色彩组合的多样性、数据特征的复杂性(如空间频率、数据连续性等)、目标任务的动态性以及显示设备的差异性，导致人们对于色彩设计合理性的评估存在差异，因此，难以实现通用于任意可视化应用场景的色彩设计方案。以二维标量场数据为例，灰度色彩表更适合梯度感知任务(如图 1a 红色虚线框)，彩虹色彩表更适合数值读取任务(如图 1a 黄色实线框)；另外，如图 1b 所示，在传统显示器上适合表达数据差异的红蓝色彩表，在增强现实环境下由于与上下文环境存在色彩混合，导致用户无法清晰地观察数值分布。因此，为了适应不同应用场景的特定需求，设计师往往需要经历低效耗时的手动调节过程，其精心设计的色彩方案却难以拓展



a. 灰度色彩表与彩虹色彩表在不同任务的表现差异



b. 红蓝色彩表编码数据及其在 AR 模拟环境中的效果

图 1 不同数据、任务及设备下的色彩感知差异示例

至其他应用场景，且设计过程中易产生色彩的误用、滥用等问题(如广受科学家欢迎的彩虹色彩表易引起错误的模式感知^[1-2]，降低了数据分析效率)。

为了提高色彩设计的效率，研究人员探索了一系列色彩设计准则及其量化表达方法，如色彩的感知一致性(perceptual uniformity)以及可辨别性(discriminability)，为色彩设计提供了理论指导。更进一步地，为了提高色彩设计的智能处理与分析能力，一系列基于数据、任务、用户、设备等可视化应用属性的自动色彩设计方法得以提出。色彩设计的核心与难点在于，如何评价色彩设计方案的合理性；然而，不同可视化应用环境存在不同的度量标准。近年来兴起的数据可视化评估理论^[3]为探索色彩设计有效性与合理性提供了思路，即色彩设计表达为在数据、用户、任务、设备等可视化应用环境下的最优决策。基于该思路，通过构建评估模型，可利用机器智能方法提高色彩设计的精准性以及合理性，这一思路在近年来的研究中逐渐得到体现。

本文将回顾数据可视化领域色彩设计相关工作，据我们所知，目前没有中文综述文章总结数据可视化中的色彩设计研究工作，相关英文综述论文可参考 Silva 等^[4]和 Zhou 等^[5]所探讨的内容。其中，Silva 等^[4]侧重于描述色彩设计准则以及色彩设计工具；Zhou 等^[5]不仅回顾了色彩空间基础理论、介绍了色彩表生成技术，还总结了一套色彩研究的分类图以帮助读者快速根据需求查阅相关文献。与上述工作不同，本文以“数据、任务、用户和设备等可视化应用环境驱动智能色彩设计”为核心观点，从数据、任务、用户和设备等角度，介绍了色彩设计准则及其量化表达、自动色彩设计和自适应色彩设计等方向的相关问题及研究成果，论述了现有色彩研究对于探索智能化色彩设计的意义与启发，并探讨了色彩计算领域未来发展趋势。

1 色彩设计准则及其量化表达

1.1 色彩设计准则

面向数据可视化的色彩设计准则多为关注于色彩自身属性的定性经验法则，如色彩的可辨别性(discriminability)、有序性(ordering)、均匀性(uniformity)、光滑性(smoothness)、视觉重要性(visual importance)和美观性(aesthetically pleasing)等^[6]。Sloan 等^[7]提出应用于图形化显示的色彩应该考虑色彩之间的可辨别性以及有序性，即不同的

色彩易于被用户区分且色彩分布符合数值大小顺序规律。Meyer 等^[8]强调在色彩设计过程中, 应在色彩空间(如 CIELAB)选取符合用户感知均匀性的色彩表达数据, 试图保证用户所感知到的相邻色彩间距与等间距数值间的差异一致。Silva 等^[9]强调了色彩的光滑性, 即表达连续数值的色彩应该保持视觉连续性, 避免出现色彩感知的突变。事实上, 被科学家广泛采用的彩虹色彩表(rainbow colormap)由于未保证光滑性, 易产生错误的边界, 从而误导用户^[1-2]。

除色彩自身属性对于数据可视化的影响, 研究人员逐渐认识到数据特征、分析任务等可视化应用环境对于色彩设计的重要性。Ware^[10]基于 5 种具有不同特征的色彩表(包括线性灰度色彩表、感知均匀的灰度色彩表、黑红线性色彩表、彩虹色彩表、红绿色彩表)开展心理物理实验, 探索了不同色彩设计对于用户感知数值大小以及模式结构的影响, 提出一系列具有启发性的色彩设计准则, 如色调的多样性有助于读取数据、亮度的单调性有助于观察数据结构模式以及整体分布。更进一步, Bergman 等^[11]对现有设计准则进行分类, 针对不同类型(比例数据、区间数据)、数据频率(高频以及低频数据)、分析任务(分类任务及突出区间任务)建立其与色彩设计属性(如亮度是否均匀、是否单调递增, 色调是否包含互补色, 饱和度是否为单调递增)之间的关联, 以此作为规则来指导色彩设计。

近年来, 色彩设计准则的研究呈现精细化的趋势, 所探索的任务类型更为多样、数据特征更为复杂、色彩特征更为丰富。Reda 等^[12]针对数值估计、梯度感知、模式感知等常见任务, 探索了 9 种在亮度、色调、色度上具有不同分布的连续色彩表与数据空间频率之间的相关作用关系。实验表明, 数据空间频率对于色彩感知有很大影响, 且具体效果取决于可视化任务。例如, 为了提高数值估计的准确性, 应尽可能地增大色彩设计中的色调差异, 这一过程可不需要考虑数据空间频率分布, 对于高频率数据下的梯度感知任务, 可使用高饱和度多色调(如彩虹色彩表)或者发散多色调(如冷暖色彩表)的色彩设计。Ware 等^[13]提出一种评估色彩表可区分度的实验方法, 即通过让用户点击色彩表达中空间特征消失的位置来评估局部可区分度。该实验在 3 种不同的数据频率模式下以及 9 种不同特征的色彩表下开展。结果表明, 感知一致的色彩空间(如 CIELAB)不能有效地预测局部可区分特征, 并在此基础上提出一种改进策略。此外, Das-

gupta 等^[14]围绕 39 位气候学家开展心理物理实验, 探索不同色彩表对于气候学家感知幅值差异、空间相似性等气候建模任务的差异。结果表明, 亮度单调的色彩表让用户更精确地感知幅值差异, 但是在空间比较任务方面效果较差。在定性色彩设计准则的基础上, 一系列的交互式色彩设计工具得以提出。举例而言, Bergman 等^[11]基于对现有色彩设计准则的分类提出了色彩设计工具 PRAVDAColor, 激励更多工作围绕色彩设计准则开展交互式色彩设计工作的研究^[15-17]。Samsel 等^[17]提出交互式设计工具 ColorMoves, 通过提供不同的数据统计特征(如直方图、等值线等)以及一系列具有不同色调范围、饱和度层次、亮度变化的色彩表, 帮助用户高效地为不均匀数值选取色彩设计方案, 从而更清楚地表达特定数值区域的细节特征(交互界面如图 2 所示, 其中, A 为色彩表选择功能栏; B 为可视化结果主视图; C 为数据分布直方图; D(左)为数据分区、色彩表插入、直方图数目等功能控制栏, D(右)为输出、不透明度等功能控制栏, D(中)为保存、输出等功能控制栏; E 为高亮直方图对应等值线区域)。

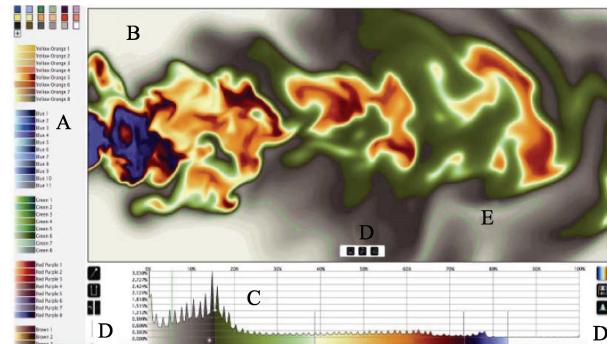


图 2 ColorMoves 界面^[17]

尽管色彩设计准则提高了数据可视化中的色彩设计效率, 但要求用户具有较强的色彩理论基础以及设计修养。因此, 如何利用机器智能建立自动的色彩设计方案, 是目前研究人员探索的重要发展方向。解决这一问题的核心在于对色彩设计合理性进行定量化表达与建模。

1.2 色彩设计量化表达

数据可视化中的色彩设计量化表达主要可以分 2 个方面的研究: 基于色彩设计准则的数学表达, 以及基于心理物理实验的色彩感知模型。

基于色彩设计准则的量化表达通常围绕可辨别性、均匀性和顺序性等色彩设计准则开展。

Bernard 等^[18]围绕查询、视觉重要性、比较等任务,针对最小可觉差异色彩(just noticeable difference colors, JND Colors)、与黑白背景的色彩距离、JND 区域大小、视觉重要性、线性感知等色彩设计准则提出了相应定量表达。其中, JND Colors 表达为二维色彩表中不重叠 JND 区域的数目;与黑白背景的色彩距离则表示为色彩表中各色彩与背景之间的 CIELAB 感知距离的最小值;JND 区域大小通过所有 JND 区域大小的标准差来计算,该数值越小,则表示色彩分布越均匀;视觉重要性则表达为采样点的视觉注意程度标准差,其中视觉注意程度表达为 CIELAB 色彩空间亮度值与色度值之间的直角三角形斜边;线性感知定义为数据空间距离及其在色彩感知空间距离的一致性。这一工作不仅提出了若干色彩准则定量表达,同时结合不同的任务探讨了定量准则的具体应用。此外, Bujack 等^[6]总结了近年来色彩设计准则相关研究,将其分成可感知的(perceptual)、可数学表达的(mathematical),以及可操作的(operational)3 种准则类型;同时建立了部分可感知准则与可数学表达准则之间的关联,如基于色彩之间的感知距离、变化速度及其极值、均值、标准差等度量,针对全局或者局部感知,建立了可辨别性、均匀性、顺序性等感知准则的数学表达;并在此基础上实现了度量色彩表质量的在线工具 ColorMeasures。然而,这一工作中所描述的定量的色彩设计准则通常围绕色彩自身属性展开,难以直接应用于数据可视化复杂应用场景中,如不均匀分布的多变量数据。

由于数据可视化中的色彩设计所编码的数据存在类型与分布多样的特点,且不同用户之间存在色彩感知差异,仅考虑色彩设计中各颜色之间的关系并不能模拟真实可视化应用场景中的用户色彩感知。因此,为了探索更准确的可视化色彩设计准则,研究人员针对不同可视化应用环境提出了一系列色彩感知模型。其中, Heer 等^[19]为了提高数据可视化中色彩设计的交流与表达效率,通过用户调查实验采集用户对于不同色彩的命名,并基于所采集的 10 万组数据,利用多项式概率分布建立色彩与其命名(color-name)之间的关系模型。在此基础上,他们进一步定义色彩命名的可信度以及色彩命名距离度量矩阵,并通过色彩选择、图像编辑、调色板设计等应用表明该模型的有效性。更进一步地, Szafir^[20]针对散点图、柱状图和线图等具有不同形状属性的可视化表达(如图 3 所示),利用众包平台采集不同类型可视化表达之间的色

彩感知差异数据,指出在不同表达的色彩感知差异与编码形状相关,如在细长柱状图下的色彩差异比相同大小的散点图更容易辨别。基于上述实验数据, Szafir^[20]对经典色彩计算 CIELAB ΔE 中各项参数进行形状大小相关的标准化处理,从而构建出针对不同可视化表达的色彩差异概率模型。这一工作将可视化应用环境与色彩感知相结合,为探索可视化中定量色彩感知模型提供了新思路。

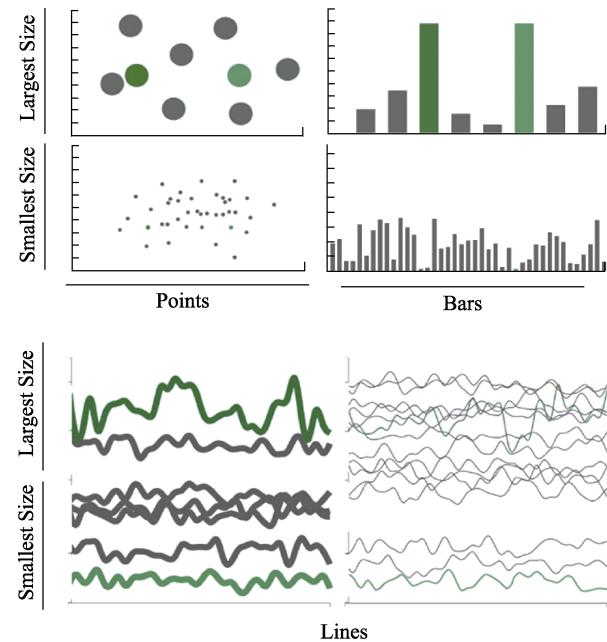


图 3 色彩在不同可视化表达的感知差异^[20]

2 自动色彩设计

自动色彩设计是提高数据可视化设计效率、节约设计成本的一个重要途径。本节将回顾数据可视化、人机交互以及计算机图形等领域的自动色彩设计方法,结合数据、任务、用户、设备等可视化应用环境,从色彩选择、色彩优化、色彩生成、色彩迁移等 4 个方面探讨研究进展。

2.1 色彩选择

色彩选择指从数量众多的色彩设计中高效地选择符合需求的方案。为了提高数据可视化的设计效率,设计师或研究人员手动创建了一系列具有不同特征的色彩表供用户选择,如亮度线性变化的灰度色彩表(grayscale)、亮度非线性变化且色调变化强烈的彩虹色彩表(rainbow),以及两端发散的冷暖色彩表(coolwarm)。随着色彩表数量的增加,如 VTK(the visualization Toolkit)中包含 62 种

默认色彩表、Python Matplotlib 中包含 164 种默认色彩表, 用户从默认色彩设计中选择合理的色彩方案愈加困难, 这一问题促进了智能色彩选择算法的发展。

PRAVDAColor^[11]是最为经典的交互式色彩选择工具, 它基于数据及任务相关的规则自动地为用户提供备选色彩设计方案。在此基础上, ColorBrewer^[15]提供结合用户指定的编码需求(如数据类别、数据属性、输出设备等)自动提供备选色彩设计策略, 如图 4 所示。上述 2 种色彩选择工具的共

同缺点在于分类粒度粗, 且无法根据用户自定义来选择满足需求的色彩表。此外, Gramazio 等^[16]提出了色彩选择工具 Colorgorical, 其核心在于可根据用户指定的可辨别性以及喜爱程度, 半随机迭代地在 CIELAB 色彩空间自动选取色彩, 以生成用于表达分类数据的离散色彩表。这一过程依赖于色彩感知距离、语义名称差异、语义名称独特性, 以及受喜爱程度等约束条件开展。然而, Colorgorical 主要针对分类任务以及离散色彩表, 难以直接推广至其他任务类型以及连续色彩表。

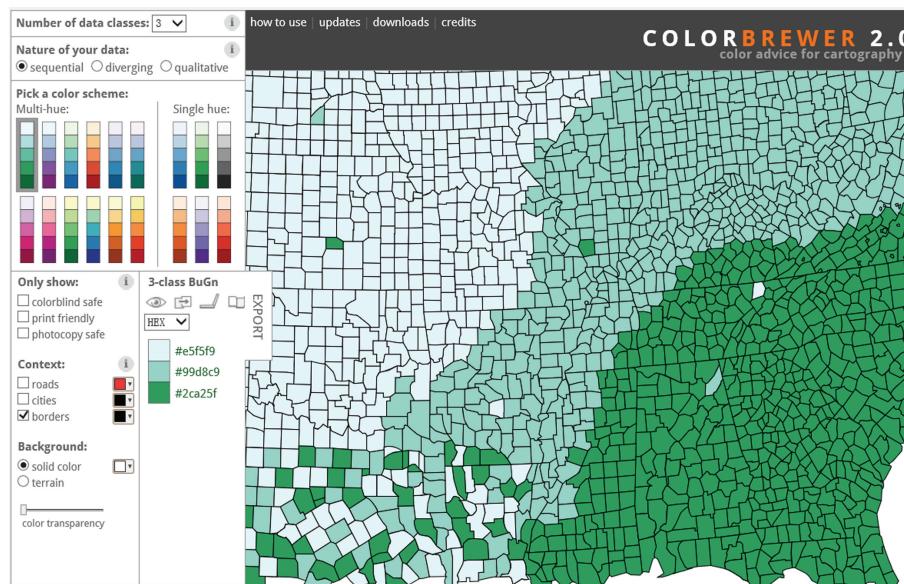


图 4 ColorBrewer 在线色彩选择工具交互界面^①

部分研究人员围绕图形设计中的色彩选择开展了一系列色彩选择交互研究。例如, Shugrina 等^[21]针对创意设计过程探索了一种支持混色交互的色彩选择工具, 通过邀请艺术家采用 Adobe Sketch 进行艺术创作并反馈相应色彩工具的优缺点, 提出高效地选择色彩、提供可兼容色彩混合、快速选择色彩组合、历史色彩回顾与删除和语义色域等需求, 为创意设计中的色彩选择工具提供了开发依据。更进一步地, 为了解决图形设计中色彩选择通常只支持小规模色彩样本选取的问题, Shugrina 等^[22]提出了一种基于色彩样本的连续色彩生成策略, 并构建了一种支持色彩样本选取、连续/离散梯度以及空间色彩混合等功能的色彩选择工具 ColorBuilder。然而, 上述工作主要针对图形艺术创作开展, 其交互有效性在数据可视化艺术作品设计过程中的影响有待进一步探索。

总体而言, 色彩选择工具提高了非专业用户

选择色彩表的效率, 但其智能化功能受限于色彩表选取过程中所依赖的粗粒度色彩设计准则。用户交互是解决该问题的一个思路。此外, 自动融入数据、任务、设备等可视化应用因素, 以追求更为智能化的处理方法仍然是需要探索的方向。

2.2 色彩优化

色彩优化过程旨在对默认色彩表或其所编码得到的可视化图像进行计算, 使得调整后的色彩表达可在某种度量标准下达到更优。

在针对色彩表的优化中, 部分方法依赖于色彩自身属性指导优化过程, 以使得优化后的色彩满足诸如亮度递增性、感知均匀性等可量化的色彩设计准则。例如, Sisneros 等^[23]以亮度单调递增、与原色彩表差异较小为约束, 不断通过校正色度分量(LUV 空间)优化色彩分布; Nuñez 等^[24]面向色视力缺乏的色盲患者, 基于感知均匀性原则设计自动生成色彩表。然而, 可视化表达的优劣不仅与色

^① <https://colorbrewer2.org/>

彩编码方式相关, 更与数据、任务、用户、显示设备等应用环境有着密不可分的联系^[25]. 因此, 越来越多的方法将可视化应用环境引入色彩优化的自动分析过程中.

在数据与任务驱动的色彩优化算法方面, Tominski 等^[26]曾提出直方图均衡化法, 以数据的直方图分布为依据调整色彩分布, 将更多的色彩用于表达高密度数值区域, 从而帮助用户更高效地读取数值. Thompson 等^[27]针对大规模数据提出了自动色彩优化方法, 其核心在于利用数据概率密度识别显著数值, 基于数据累积分布密度建立不同显著区间数据与色彩之间的映射. Lee 等^[28]提出一种面向可见性任务的离散色彩表优化方法, 通过计算每个样本数据的视觉显著程度计算类分离度, 在此基础上定义能量优化方程迭代计算更有助于区分类别数据的色彩分布. 该方法主要用于地图数据的可视化的数值读取, 在连续数据的应用上存在计算时长的问题. 上述方法数据及任务驱动的方法未考虑色彩属性, 可能导致优化后的色彩表达与默认色彩表存在较大差异, 不符合用户的使用习惯. 更进一步地, Zeng 等^[29]通过与领域专家深入分析与交流, 总结科学可视化应用中的 3 点主要需求——保持空间差异、色彩保真以及可交互, 并基于这 3 点需求构建面向数值变化任务的色彩优化方法, 结果如图 5 所示, 虚线框图内标明了两者在特征表达上的差异; 该方法将数据中边界

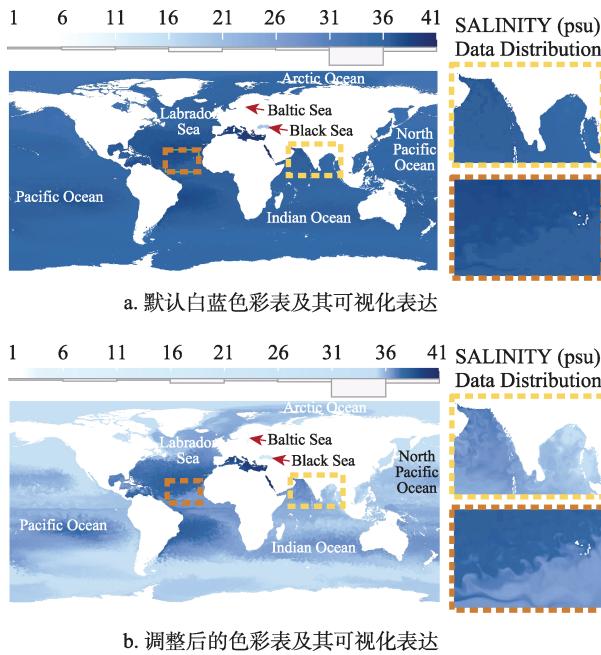


图 5 面向数值变化任务的色彩优化方法
在海洋盐度数据上的可视化结果^[29]

特征以及色彩表保真特征作为约束子, 通过求解非线性约束能量方程得到最优解, 并提供交互方式供用户探索不同区域数据特征.

部分研究强调了设备对于用户感知的影响, 并将其应用于色彩优化过程. Gresh^[30]探索性地基于物理显示设备迭代优化色彩分布, 基于色彩感知经验实验, 建立用户在特定物理显示设备(如 IBM P201 阴极射线管显示器)下针对某一色彩表(如彩虹色彩表)的色彩感知函数. 该函数描述了色彩表参数位置与最小可觉偏移之间的关联, 通过调整色彩分布达到该曲线的分布均匀化是本文优化色彩的基本目标. 然而, 上述过程依赖于不断重复且设备相关的用户感知实验, 整个过程耗时长且难以拓展至普通应用环境. Waldin 等^[31]采用色彩排序实验测量用户在特定显示器上区分二维彩表中各色彩的能力, 利用主成分分析度量用户的色彩感知错误, 并在感知错误的基础上基于质点弹簧模型对二维色彩表进行形变, 使得变化后的色彩表更适用于特定的用户及显示器. 也有研究将色彩优化应用于显示器节能过程中. 例如, 浙江财经大学 Chen 等^[32]针对体绘制可视化探索了面向有机发光二极管显示器能量节约的色彩优化策略, 该方法基于一种亮度相关的能量估计模型预估显示器能量消耗, 融合可区分性等色彩准则约束而生成最优色彩方案.

此外, 部分研究人员探索了基于可视化编码图像的色彩优化方法. Elmqvist 等^[33]针对用户指定区域的样本透镜(lens)动态优化色彩表达, 建立层级表达模式(根节点为整个可视化图像, 叶子节点分别为各局部透镜), 并根据优化过程中是否仅考虑局部直方图、是否将色彩表达迁移至所有层级等因素构建了 4 种不同的透镜, 帮助用户高效地进行交互式探索. Zhou 等^[34]针对高动态范围标量场数据提出一种融合色调映射以及高光模拟的可视化表达色彩增强方法, 其色调映射能够在保持原始数值与结构的情况下压缩数据范围, 高光模拟计算的引入有助于提高用户感知局部区域的注意力程度. 在此基础上, 他们提出了 3 种魔术透镜模式帮助用户探索数据, 包括如图 6 所示显像模式、像素高亮模式以及对比度增强模式. 然而, 该方法改变了原始数值, 且在对比度增强模式下不能完全保证数值与色彩表的一一映射关系, 一定程度阻碍了用户直观地读取数值. 这一缺点在其他基于可视化编码图像色彩优化方法亦存在, 如何在提高色彩表达效率的情况下保持色彩感知的准确性, 是该方向可进一步探究的课题.

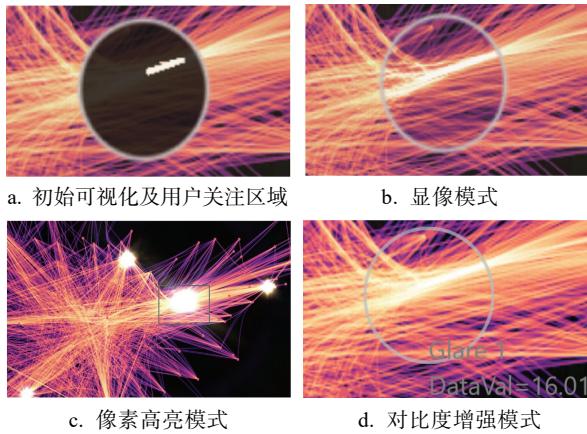


图 6 融合色调映射及高光模拟的可视化色彩增强结果^[34]

2.3 色彩生成

色彩生成旨在根据用户需求或色彩设计准则自动构造离散或连续色彩表的过程。现有色彩表生成方法研究大多是基于可定量表达的色彩设计准则开展的。其中, Levkowitz 等^[35]在生成色彩表的过程中考虑了有序性、感知一致性等易于构建数学表达的色彩设计度量准则; Wijffelaars 等^[36]提出一种色彩设计的参数化表达方式, 将色彩表达为色调、饱和度、亮度、色彩数目、冷暖程度以及对比度等参数形式, 通过定义在色彩空间具有均匀色彩距离、有限数目色彩的连续路径生成离散色彩表, 可有效地提高色彩生成效率。

近年来, 越来越多的研究探索如何将任务需求作为色彩生成的依据。Mittelstädt 等^[37]提出一种基于混合任务的色彩生成方法及交互式系统——ColorCAT, 该系统通过对不同元任务(定位、识别以及比较)组合下的视觉重要程度、色彩可区分度、感知一致性等定量色彩度量准则建模, 生成符合用户给定数据及任务的色彩表, 所生成的色彩表同时适用于色盲用户, 并在此基础上提供了文化、艺术性等相关的色彩生成交互式工具。Smart 等^[38]提出一种符合设计专家设计特点的色彩表自动生成方法, 其首先采集 222 位专家设计的色彩表, 并对该表进行基于曲线拟合的标准化处理, 通过聚类算法(贝叶斯或 K-means 聚类)构建专业色彩曲线模型。对于用户给定的单个种子色彩, 可根据该曲线模型插值生成符合专业设计风格的色彩设计方案。该方法由于采用聚类方法对专业设计色彩表进行统一色彩表达, 在提取色彩表共性的过程中丢失了专业设计色彩表的多样性特点(如图 7 所示, 用户选择单个种子色彩, 算法可自动生成符合种子色彩的离散色彩表)。

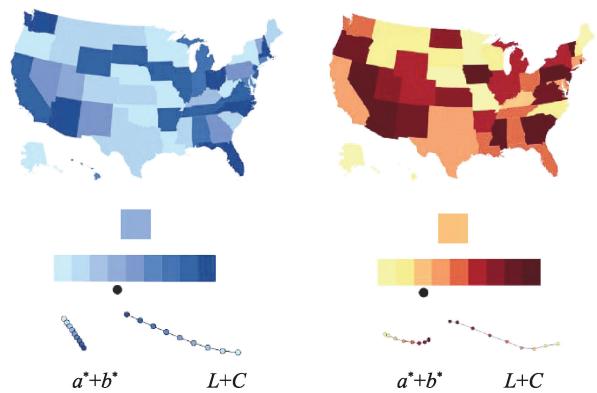


图 7 色彩生成方法的结果^[38]

此外, 部分研究借助于图像中丰富的色彩表达生成语义相关的色彩板^[39-40]。其中, Lin 等^[39]对用户感知色彩主题的过程进行建模, 通过 AMT 众包平台采集一系列用户指定的可表征图像的离散色彩, 并基于 LASSO 回归线性模型构造离散色彩表的评分依据, 建立了显著性、色彩覆盖比率、色彩多样性、色彩纯度, 以及色彩可命名程度、聚类统计等特征与色彩主题感知之间的关系。Nguyen 等^[40]提出一种类别图像数据(如香蕉图像)驱动的色彩选择空间构建方法, 通过采集同一类别的彩色图像并统计各色彩的密度分布, 利用主成分分析、多维缩放和自组织特征映射等方法分别对高维密度分布降维, 从而得到语义相关的低维色彩表达, 并进一步将其作为色彩选择器, 帮助用户精确调节或选择色彩。该方法缺陷在于, 其语义信息强烈依赖于图像的类别, 而真实设计过程中图形元素存在多样性。

尽管上述色彩生成方法为设计师、领域用户生成符合美观感知的色彩表提供了智能化工具, 但是在科学领域实际的应用中, 为了更合理地完成任务需求, 大部分色彩生成工作仍然由专业设计师与科学家联合完成, 该过程需要双方花费大量的时间沟通目标任务^[41-42]。因此, 在智能化色彩生成的过程中进一步强调数据、任务的作用, 精细化数据特征、任务表达等对于色彩生成效果的影响, 将对智能色彩生成的实际应用有重要作用。

2.4 色彩迁移

色彩迁移旨在根据目标图像(或其他色彩需求)改变输入图像中的色彩分布, 使得色彩迁移后的图像仍然能够保持原始图像中的空间结构或语义信息。这一问题在计算机图形学与图像处理中得到广泛的研究^[43-44]。例如, Eynard 等^[43]提出基于图像的拉普拉斯矩阵度量结构信息, 通过拉普拉斯矩阵可交换性建立色彩质量评估准则, 使得在

色彩迁移的过程中尽可能地保证变化前后图像的拉普拉斯矩阵特征向量的一致性。该方法可在针对色盲用户的色彩迁移过程中有效地保持信息。此外, He 等^[44]基于用户给定的参考色彩, 利用深度学习网络获取内容语义相关的色彩迁移效果, 其核心在于利用亮度通道上的图像分类网络提取语义关联信息以及双向映射函数预估内容语义区域的迁移色彩。

上述色彩迁移过程追求图像结构及语义信息的一致性, 未考虑色彩与数值之间的关联。与传统色彩迁移不同, 数据可视化中的色彩迁移需要尽可能地保持原始数据的准确性。为了达到这一目标, 一个典型的思路是从可视化表达中准确地提取数值, 再进行色彩映射。Poco 等^[45]提出了一种半自动的从位图中提取数值并重新映射色彩的方法, 它以人工标记可视化图像中离散或者连续的色彩图例为基础, 训练深度学习网络识别色彩图例的属性, 在此基础上采用不同方法提取色彩数值, 从而映射新的色彩, 如图 8 所示。该方法是可视化领域关于可视化表达提取以及色彩迁移的开创性探索; 但其不可避免地依赖于色彩图例的识别精度, 且结果受限于原始色彩表达的精确性, 在高动态范围数据的处理中可能丢失大量原始数据信息。

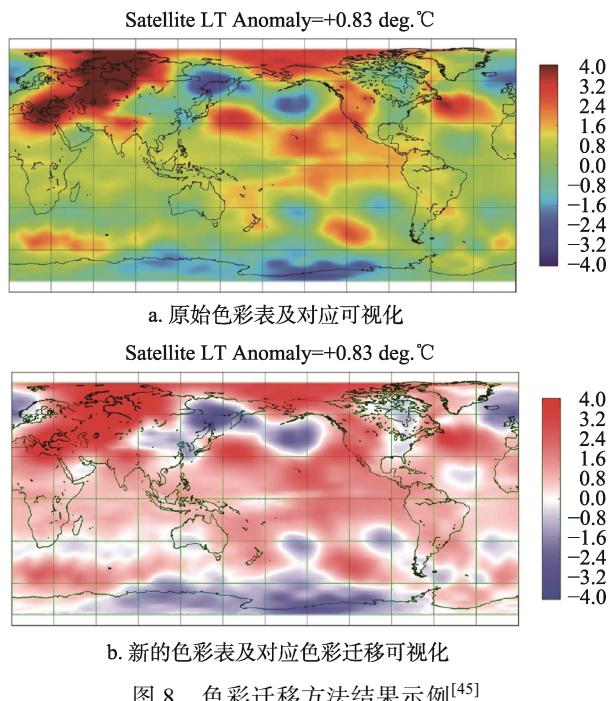


图 8 色彩迁移方法结果示例^[45]

目前, 面向数据可视化的色彩迁移未得到足够重视。事实上, 该领域还有许多值得探索的课题, 例如, 如何在没有色彩图例的情况下高效、准

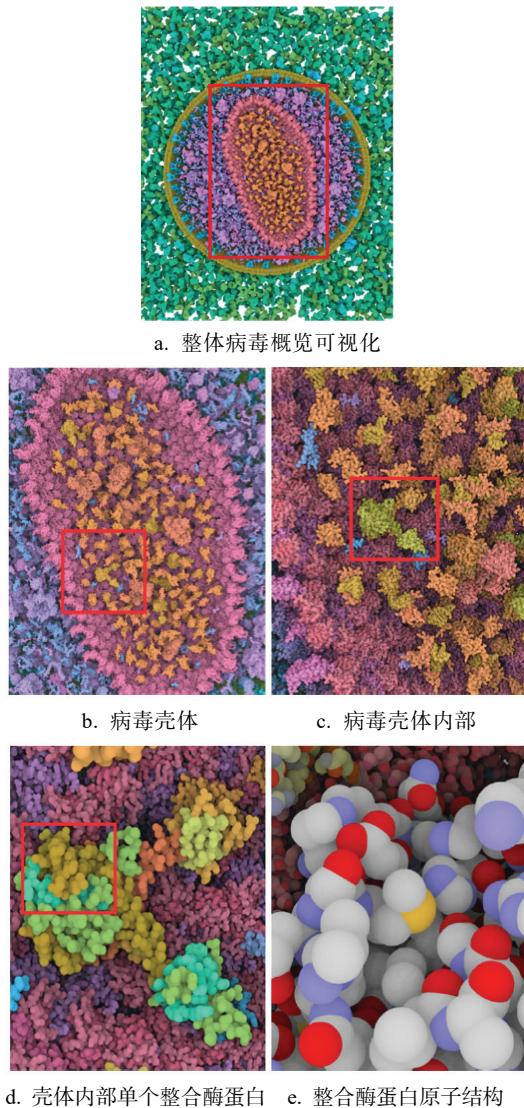
确地进行色彩迁移, 如何处理多变量数据色彩迁移等。

3 自适应色彩设计

自适应色彩设计指根据实际应用场景的动态多变性设计色彩或更新色彩方案。其核心问题在于如何在不同视角、不同时序、不同尺度的情景下保证色彩动态变化的感知连贯性和语义一致性, 以及如何在不同方案中自适应分配色彩, 以满足任务需求(如数值可区分度)。动态变化场景下的自适应色彩设计相关研究近年来逐渐受到关注, 多针对动态多尺度数据开展。具体而言, Waldin 等^[46]针对分子可视化中的多尺度结构(如分子、次生结构、氨基酸、原子等), 动态地调整色彩设计方案, 使得用户能够清晰地区分当前尺度的结构信息。其多尺度处理的核心思想是, 在高尺度基于力导向分析为各蛋白质结构分配具有不同色相的色彩; 在低尺度下则基于可见性更新色彩饱和度提高当前视角下的色彩可区分度, 并利用插值算法建立不同交互视角之间的色彩过渡。该方法可能导致不同尺度内的色彩相同, 且主要针对分子可视化开展。在此基础上, Waldin 等^[47]进一步提出一种动态自适应色彩设计机制, 使得色彩设计可反映层次结构且符合色彩语义相关性及一致性, 结果如图 9 所示(红色框标记了缩放区域)。与上一工作的主要差别在于, 该方法利用亮度信息编码不同尺度的累加数值建立层次结构关联。

现有自适应色彩分配方案的研究, 多围绕离散色彩设计开展。其中, Setlur 等^[48]针对具有强色彩关联的类别数据(如玉米、胡萝卜等蔬菜水果类别), 基于探索语言与色彩之间的关联关系, 并基于该关系自适应地设计相应类别数据的色彩分配方案。此外, Wang 等^[49]提出一种感知驱动的散点图色彩分配方法, 其引入多类散点数据的空间分布关系、密度、点群间的重叠程度, 以及背景颜色等因素, 构建优化方程, 并基于遗传算法优化多类散点图色彩分配方案。

动态变化场景下的自适应色彩设计方法多为基于色相轮的启发式算法, 在交互缩放时仅根据亮度调整色彩可能导致色彩变化显著而误导用户。另外, 多着色方案的自适应色彩设计方法主要针对类别数据的可区分性任务开展, 对于其他可视化表达(如饼图、平行坐标等), 以及任务情景(如趋势分析)的有效性仍需进一步探索。

图9 动态多尺度色彩映射方法结果^[47]

4 总结与展望

本文从数据、任务、用户、设备等角度分析了色彩设计准则及其量化表达、自动色彩设计和自适应色彩设计的前沿工作。综合国内外研究现状及发展分析,现有色彩设计过程缺乏定量色彩评估模型指导,且缺乏统一的自适应计算框架,因此难以避免在实际色彩设计过程中所存在的低效耗时、难以拓展以及色彩误用和滥用等问题。众所周知,数据可视化中的色彩感知是一个主客观条件相互博弈的过程,不仅与用户主观感知有关,更与数据、任务和设备等客观的可视化应用环境存在不可分离的关系;这一博弈过程在色彩设计相关研究中日渐重视。因此,本文认为,未来可视化色彩计算领域的发展大方向主要为以下3个方面:

(1) 依据数据、任务和设备等因素探索精细化的用户色彩感知模型与色彩设计准则。目前色彩感知模型与色彩设计准则往往依赖于有限分类的色彩表、任务以及数据类型开展研究,未建立各因素数学表达之间的关联。以元任务的抽象表达^[3]、统计数值特征^[50]为基础,探索其与色彩分量之间的感知关联,是构建精细化色彩感知模型与色彩设计准则的重要思路。

(2) 利用机器智能建立任务有效的自动色彩设计。该问题的核心在于建立色彩有效性评估机制,由于不同领域数据、任务、显示设备存在差异而导致自动色彩设计算法有所区别,没有统一的自动色彩设计框架。因此,探索面向不同领域的共性评估准则,基于共性色彩评估模型研究面向领域个性需求的统一自动色彩计算框架与方法具有重要意义。

(3) 自适应地更新色彩设计以高效引导交互。现有自适应色彩设计多强调多尺度动态过程的感知一致性,未来在时序数据的动态色彩设计可进一步探索;此外,动态色彩设计的目标在于帮助用户提高数据探索效率,将视觉注意力机制引入色彩设计过程有助于提高交互效率。

参考文献(References):

- [1] Rogowitz B E, Treinish L A. Data visualization: the end of the rainbow[J]. IEEE Spectrum, 1998, 35(12): 52-59
- [2] Borland D, Taylor R M. Rainbow color map (still) considered harmful[J]. IEEE Computer Graphics and Applications, 2007, 27(2): 14-17
- [3] Brehmer M, Munzner T. A multi-level typology of abstract visualization tasks[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(12): 2376-2385
- [4] Silva S, Santos B S, Madeira J. Using color in visualization: a survey[J]. Computers & Graphics, 2011, 35(2): 320-330
- [5] Zhou L, Hansen C D. A survey of colormaps in visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(8): 2051-2069
- [6] Bujack R, Turton T L, Samsel F, et al. The good, the bad, and the ugly: a theoretical framework for the assessment of continuous colormaps[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1): 923-933
- [7] Sloan K R, Brown C M. Color map techniques[J]. Computer Graphics and Image Processing, 1979, 10(4): 297-317
- [8] Meyer G W, Greenberg D P. Perceptual color spaces for computer graphics[J]. ACM SIGGRAPH Computer Graphics, 1980, 14(3): 254-261
- [9] Silva S, Madeira J, Santos B S. There is more to color scales

- than meets the eye: a review on the use of color in visualization[C] //Proceedings of the 11th International Conference Information Visualization. Los Alamitos: IEEE Computer Society Press, 2007: 943-950
- [10] Ware C. Color sequences for univariate maps: theory, experiments and principles[J]. IEEE Computer Graphics and Applications, 1988, 8(5): 41-49
- [11] Bergman L D, Rogowitz B E, Treinish L A. A rule-based tool for assisting colormap selection[C] //Proceedings of the 6th Conference on Visualization. Los Alamitos: IEEE Computer Society Press, 1995: 118-125
- [12] Reda K, Nalawade P, Ansah-Koi K. Graphical perception of continuous quantitative maps: the effects of spatial frequency and colormap design[C] //Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2018: Article No.272
- [13] Ware C, Turton T L, Bujack R, et al. Measuring and modeling the feature detection threshold functions of colormaps[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(9): 2777-2790
- [14] Dasgupta A, Poco J, Rogowitz B, et al. The effect of color scales on climate scientists' objective and subjective performance in spatial data analysis tasks[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(3): 1577-1591
- [15] Harrower M, Brewer C A. ColorBrewer.org: an online tool for selecting colour schemes for maps[J]. The Cartographic Journal, 2003, 40(1): 27-37
- [16] Gramazio C C, Laidlaw D H, Schloss K B. Colorgorical: creating discriminable and preferable color palettes for information visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1): 521-530
- [17] Samsel F, Klaassen S, Rogers D H. ColorMoves: real-time Interactive colormap construction for scientific visualization[J]. IEEE Computer Graphics and Applications, 2018, 38(1): 20-29
- [18] Bernard J, Steiger M, Mittelstädt S, et al. A survey and task-based quality assessment of static 2D colormaps[C] //Proceedings of SPIE. Bellingham: Society of Photo-Optical Instrumentation Engineers, 2015, 93970: 93970M
- [19] Heer J, Stone M. Color naming models for color selection, image editing and palette design[C] //Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2012: 1007-1016
- [20] Szafir D A. Modeling color difference for visualization design[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1): 392-401
- [21] Shugrina M, Lu J W, Diverdi S. Playful palette: an interactive parametric color mixer for artists[J]. ACM Transactions on Graphics, 2017, 36(4): Article No.61
- [22] Shugrina M, Zhang W J, Chevalier F, et al. Color builder: a direct manipulation interface for versatile color theme authoring[C] //Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2019: Article No.456
- [23] Sisneros R, Raji M, van Moer M W, et al. Chasing rainbows: a color theoretic framework for improving and preserving bad colormaps[C] //Proceedings of International Symposium on Visual Computing. Heidelberg: Springer, 2016: 391-402
- [24] Nuñez J R, Anderton C R, Renslow R S. Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data[J]. PLoS ONE 2018, 13(7): e0199239
- [25] Munzner T. Visualization analysis and design[M]. Boca Raton: A K Peters/CRC Press, 2014
- [26] Tominski C, Fuchs G, Schumann H. Task-driven color coding [C] //Proceedings of the 12th International Conference Information Visualisation. Los Alamitos: IEEE Computer Society Press, 2008: 373-380
- [27] Thompson D, Bennett J, Seshadri C, et al. A provably-robust sampling method for generating colormaps of large data[C] //Proceedings of the IEEE Symposium on Large-Scale Data Analysis and Visualization. Los Alamitos: IEEE Computer Society Press, 2013: 77-84
- [28] Lee S, Sips M, Seidel H P. Perceptually driven visibility optimization for categorical data visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(10): 1746-1757
- [29] Zeng Q, Wang Y Q, Zhang J, et al. Data-driven colormap optimization for 2D scalar field visualization[C] //Proceedings of the IEEE Visualization Conference. Los Alamitos: IEEE Computer Society Press, 2019: 266-270
- [30] Gresh D L. Self-corrected perceptual colormaps[R]. Technical report, IBM, 2008
- [31] Waldin N, Bernhard M, Rautek P, et al. Personalized 2D color maps[J]. Computers & Graphics, 2016, 59(C): 143-150
- [32] Chen W F, Chen W, Chen H D, et al. An energy-saving color scheme for direct volume rendering[J]. Computers & Graphics, 2016, 54: 57-64
- [33] Elmqvist N, Dragicevic P, Fekete J D. Color lens: adaptive color scale optimization for visual exploration[J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 17(6): 795-807
- [34] Zhou L, Rivinius M, Johnson C R, et al. Photographic high-dynamic-range scalar visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(6): 2156-2167
- [35] Levkowitz H, Herman G T. Color scales for image data[J]. IEEE Computer Graphics and Applications, 1992, 12(1): 72-80
- [36] Wijffelaars M, Vliegen R, van Wijk J J, et al. Generating color palettes using intuitive parameters[J]. Computer Graphics Forum, 2008, 27(3): 743-750
- [37] Mittelstädt S, Jäckle D, Stoffel F, et al. ColorCAT: guided design of colormaps for combined analysis tasks[C] //Proceedings of Eurographics Conference on Visualization. Aire-la-Ville: Eurographics Association Press, 2015: 115-119
- [38] Smart S, Wu K K, Szafir D A. Color crafting: automating the construction of designer quality color ramps[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(1): 1215-1225
- [39] Lin S, Hanrahan P. Modeling how people extract color themes from images[C] //Proceedings of the SIGCHI Conference on

- Human Factors in Computing Systems. New York: Association for Computing Machinery, 2013: 3101-3110
- [40] Nguyen C H, Ritschel T, Seidel H P. Data-driven color manifolds[J]. ACM Transactions on Graphics, 2015, 34(2): Article No.20
- [41] Samsel F, Petersen M, Geld T, et al. Colormaps that improve perception of high-resolution ocean data[C] //Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2015: 703-710
- [42] Samsel F, Turton T L, Wolfram P, et al. Intuitive colormaps for environmental visualization[C] //Proceedings of the Workshop on Visualisation in Environmental Sciences. Aire-la-Ville: Eurographics Association Press, 2017: 55-59
- [43] Eynard D, Kovnatsky A, Bronstein M M. Laplacian colormaps: a framework for structure-preserving color transformations[J]. Computer Graphics Forum, 2014, 33(2): 215-224
- [44] He M M, Chen D D, Liao J, et al. Deep exemplar-based colorization[J]. ACM Transactions on Graphics, 2018, 37(4): Article No.47
- [45] Poco J, Mayhua A, Heer J. Extracting and retargeting color mappings from bitmap images of visualizations[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1): 637-646
- [46] Waldin N, Muzic M L, Waldner M, et al. Chameleon: dynamic color mapping for multi-scale structural biology models[C] //Proceedings of Eurographics Workshop on Visual Computing for Biology and Medicine. Aire-la-Ville: Eurographics Association Press, 2016: 11-20
- [47] Waldin N, Waldner M, Muzic M L, et al. Cuttlefish: color mapping for dynamic multi-scale visualizations[J]. Computer Graphics Forum, 2019, 38(6): 150-164
- [48] Setlur V, Stone M C. A linguistic approach to categorical color assignment for data visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1): 698-707
- [49] Wang Y H, Chen X, Ge T, et al. Optimizing color assignment for perception of class separability in multiclass scatterplots[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(1): 820-829
- [50] Hu K, Bakker M A, Li S, et al. VizML: a machine learning approach to visualization recommendation[C] //Proceedings of the CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2019: Article No.128