

# 陈昕

山东大学/计算机科学与技术/博士 · (+86) 16653237032 · chenxin199634@gmail.com

## 个人简历

- 陈昕, 2014-2024 年在山东大学完成本、硕、博的学习和研究生涯, 其中 2018 年获得保送资格, 预计于 2024 年 12 月拿到博士学位, 已获得山东大学优秀博士毕业生资格。我的研究方向是大规模数据的高效可视化与交互式分析, 并深入探索了颜色映射的自动分配与生成技术。读博期间, 我累计发表 CCF-A 类学术论文 5 篇, 其中一作 4 篇。另外, 在我的导师汪云海教授的指导下, 我带领硕士学弟们一起, 完成了山东大学首篇 ACM SIGMOD 论文的发表, 获得了学校的特别嘉奖。
- 在研究生生涯中, 我作为实验室的核心成员参与了多项国家自然科学基金和重点研发计划项目, 包括“基于云平台的复杂场景下群体及遮挡三维信息获取与分析”、“任务驱动的可视化方法自动选择”与“三元空间大数据的跨域可视分析基础理论与方法”等。主导研发了基于可视空间建模的多类别与渐进式/流式采样技术、基于层次建模的时序数据交互式分析技术、以及基于数据分布的颜色分配优化技术等。项目的应用成果顺利获得 2 项国家专利授权, 在应用领域产生了显著的实用价值, 这些成果支撑我获得山东省研究生优秀学术成果奖。
- 在参与大型科研项目过程中, 我擅长将实际问题抽象为数学问题, 并运用先进的数学建模知识和计算机技术来构建解决方案, 并结合定制化的数据结构与算法实现高效解决方案。项目研究成果被 IEEE TVCG、ACM SIGMOD 等数据挖掘、可视化相关的国际顶级期刊与会议收录。读博期间, 我连续五年荣获山东大学一等奖学金, 并获得了山东省研究生优秀成果奖及乾云社会奖学金等多项荣誉。

## 教育背景 & 个人荣誉

山东大学, 计算机科学与技术, 在读博士研究生 2018.9 - 2024.12

- 山东大学研究生一等学业奖学金 (2018-2022, 连续 5 年)
- 山东省研究生优秀学术成果奖三等奖 (2019)
- 乾云社会奖学金 (2022)
- 山东大学优秀毕业生 (2024)

山东大学, 软件工程, 工学学士 2014.9 - 2018.6

- 山东大学优秀学生二等奖学金 (2016-2017, 连续 2 年)
- 国际水下机器人大赛花样游泳组冠军 (2016)
- 国家励志奖学金 (2017)

## 项目经历

任务驱动的可视化方法自动选择 (国家自然科学基金项目) 2018 - 2021

- 项目背景:** 随着可视化技术的理论研究在实际应用场景的广泛运用, 大量用户面临着如何选择恰当的可视化方式来有效展现数据特征的“选择困境”挑战。为此, 该项目聚焦于根据分析任务需求进行可视化方法分类及自动化设计, 本人作为项目骨干成员, 担任大规模数据自动可视化方向的核心工作。
- 关键技术:**
  - 针对海量分类数据提出了非均匀采样技术: 首先采用统计方法将数据映射至可视空间, 依据可视空间中的密集度特征构建 KD 树结构; 接着在类别约束下进行回溯搜索, 通过轮盘赌选择法分配类别标签并执行采样。该方法创新性地解决了面对海量数据分析所需的全局与局部特征如何同时保留的经典难题。
  - 针对渐进式与流式数据提出了动态采样技术: 将数据的密集度特征建模为多分辨率金字塔, 制定双目标分配策略来执行一过性快速采样; 在动态场景下, 依据多尺度比率关系识别出显著变化的区域进行局部更新, 解决了传统方法的采样延迟高和视觉结果不连贯的双重难题。

三元空间大数据的跨域可视分析基础理论与方法 (国家自然科学基金项目) 2022 - 2023

- 项目背景:** 物联网、云计算与移动互联网技术的发展促进了网络-物理-社会 (CPS) 三元大数据的形成, 其跨领域特性、复杂性和不确定性对人工智能等现有分析方法构成重大挑战。该项目致力于开发创新可视分析工具以增强对 CPS 大数据的理解, 本人负责设计面向海量数据的实时交互式可视分析技术。
- 关键技术:** 针对超大规模时序数据, 利用数学建模知识, 首先从理论上证明了线段聚合能保证仅使用必要的最大最小值样本即可生成准确的折线图可视化, 然后利用计算机技术, 提出了一种形似完全二叉树的有序最大最小值层次结构和能够在任意宽度屏幕上高效搜索必要样本的增量查询与剪枝策略, 接着针对这种特定的层次结构和策略, 采用软件工程理论和实践, 设计并实现了相应的系统原型, 在实际的应用场景中能保证准确性且大幅减少传输数据量, 支持缩放等多种交互式操作, 解决了现有技术响应速度慢及可重用性差的问题。

## 专业技能

---

- 编程语言: Python, C++, JavaScript, HTML/CSS, SQL, Shell
- 开发环境: NumPy, Pandas, Matplotlib, Scikit-learn, Qt, Vue.js, OpenCV, PostgreSQL, InfluxDB, Linux
- 英语等级: CET-6

## 科研成果

---

### **A Recursive Subdivision Technique for Sampling Multi-class Scatterplots** 2019

- IEEE Transactions on Visualization and Computer Graphics (SCI 1 区, 一作)
- 提出了一种非均匀多类采样技术, 基于数据到可视空间的映射关系对记录数量建模, 随后利用定制化的 KD 树结构划分可视空间, 再采用回溯搜索定位适宜节点进行采样, 在揭示各类别数据分布与保留异常值间达成了平衡, 成功解决了大规模数据导致散点图过度绘制的问题。

### **Pyramid-based Scatterplots Sampling for Progressive and Streaming Data Visualization** 2021

- IEEE Transactions on Visualization and Computer Graphics (SCI 1 区, 一作)
- 提出了一种适用于渐进式与流式数据的动态采样技术, 采用多分辨率金字塔结构对可视空间中的数据记录数量建模, 快速进行自顶向下采样与微调以同时保持数据分布与异常值。通过利用金字塔中各层级的比率关系, 在连续帧之间局部更新变化显著的区域, 有效解决了现有采样技术中存在的高延迟和视觉不连贯问题。

### **Visualization-Driven Illumination for Density Plots** 2024

- IEEE Transactions on Visualization and Computer Graphics (SCI 1 区, 一作)
- 提出了一种定制的可视化驱动光照模型, 针对密度图相关的分析任务的需求, 采用高斯差分 and 高度夸张处理生成能呈现局部细节和离群点的阴影场, 再将阴影通过 CIELAB 颜色空间的亮度通道叠加到体现全局结构的密度场上, 有效揭示中高密度区域的细节结构以及低密度区域的异常值, 同时避免出现影响密度值估计等任务的颜色失真现象。

### **OM<sup>3</sup>: An Ordered Multi-level Min-Max Representation for Interactive Progressive Visualization of Time Series** 2023

- Proceedings of the ACM on Management of Data (CCF A 类会议, 共同一作)
- 提出了一种面向十亿级别时序数据的有序最大最小值层次结构, 通过建模在大规模时序数据可视化折线图过程中的线段聚合现象, 设计了在层次结构中高效搜索必要样本的增量查询与剪枝策略, 在保证像素级准确性的同时将传输数据量减少了约 99%, 响应延迟低于 300 毫秒, 支持缩放、平移、调整大小和刷选等一系列交互式操作, 解决了现有技术中存在的响应速度慢及可重用性差的问题。

### **Optimizing Color Assignment for Perception of Class Separability in Multiclass Scatterplots** 2018

- IEEE Transactions on Visualization and Computer Graphics (SCI 1 区, 共同一作)
- 提出了一种面向多类离散数据的调色板分配技术, 将数据空间中各类别的距离与 CIELAB 颜色空间中的颜色距离度量结合形成目标函数, 再通过遗传算法优化找出最佳颜色分配方案, 帮助用户准确地感知各类的聚簇、趋势等模式, 避免了类间区分度不足造成的对数据分布的误解。

### **Palettaylor: Discriminable Colorization for Categorical Data** 2020

- IEEE Transactions on Visualization and Computer Graphics (SCI 1 区, 三作)
- 提出了一种数据驱动的为不同可视化图表自动着色的通用方法, 适用于多类散点图、折线图和直方图, 显著提升了类间的区分度。

### **Bi-Scale Density-Plot Enhancement based on Variance-Aware Filter** 2024

- 国内计算机图形学领域顶级会议 Chinagraph (SCI, 二作)
- 提出了一种双尺度密度图增强技术, 通过边缘保持的方差感知过滤器提取局部密度的细微变化, 允许用户灵活地调节细微变化的权重, 增强感兴趣的结构的可视化效果。