

地理空间数据可视化中的过滤

陆 旻, 袁晓如*

(北京大学信息科学技术学院机器感知与智能教育部重点实验室 北京 100871)
(xiaoru.yuan@pku.edu.cn)

摘要: 交通、气象等领域通过收集物体的空间位置信息(地理空间数据)来分析其潜在的运动规律。可视分析因其直观的可视化表达与便捷的交互方式, 能有效地分析地理空间数据。过滤作为一种重要的数据操作, 被广泛地应用于大规模地理空间数据的可视分析中。文中概述了地理空间数据可视化中的过滤方法及相关工作。

关键词: 地理空间数据; 可视分析; 过滤
中图法分类号: TP391

Filter in Visualization of Geospatial Data

Lu Min and Yuan Xiaoru*

(Key Laboratory of Machine Perception (Ministry of Education), School of Electronic Engineering and Computer Science, Peking University, Beijing 100871)

Abstract: Tagged with spatial positions, geospatial data is collected in many research fields, such as transportation, meteorology etc. Visualization facilitates the understanding and analysis of those geospatial data. To deal with the large scale geospatial data, filter is widely incorporated in the visual analysis process. In this paper, we summarize the filter in geospatial visualizations and the related work.

Key words: geospatial data; visual analytics; filter

随着移动采集技术的发展和普及, 人们可以对各种各样物体的移动行为进行采样, 以获得记录物体位置信息的地理空间数据。在科学研究上, 交通、气象、生态等领域通过收集这些地理空间数据来挖掘物体的运动规律, 发现异常行为现象等。如在交通领域, 科学家通过全球定位系统(global positioning system, GPS)采集车辆的移动轨迹研究交通问题, 如道路拥堵^[1-2]。同时, 伴随着个人移动设备的普及, 地理空间数据的来源不再局限于专业的移动采集设备。在日常生活中, 人们通过各式各样的社交媒体所发布的图片、文字也常常带有地理位置标记。此类社交媒体数据蕴含着人们的

社会活动^[3-4]、社会事件与空间位置之间的关联^[5]。

可视化及可视分析将数据通过直观的视觉形式展现给用户, 并提供有效的交互方式支持用户对数据进行探索^[6]。针对地理空间数据所特有的空间信息, 可视分析方法能将其空间关系直观地展现出来, 并将自动分析算法和灵活交互方式相结合, 已经成为探索地理空间数据的一种有效分析手段。目前, 在可视化地理空间数据的方向上, 研究人员做出了相当多的探索。针对这些工作, Andrienko 等^[7]提出了移动数据可视分析的理论框架。蒲剑苏等^[8]和王祖超等^[9]则针对地理空间数据中最为常见的数据类型——轨迹数据进行了相关研究。

收稿日期: 2016-04-06; 修回日期: 2016-04-12. 基金项目: 国家自然科学基金(61170204, 61232012); 国家“九七三”重点基础研究发展计划项目(2015CB352500). 陆 旻(1988—), 女, 博士研究生, 主要研究方向为轨迹数据可视分析; 袁晓如(1975—), 男, 博士, 研究员, 博士生导师, 论文通讯作者, 主要研究方向为可视化与可视分析。

可视分析遵循着 Shneiderman^[10]提出的经典的探索流程: 先总览、缩放和过滤, 再分析细节(overview first, zoom and filter, then details-on-demand). 过滤作为一种数据操作, 可帮助用户获得感兴趣的数据, 尤其当数据规模越来越大时, 过滤在可视分析中的作用更为突出. 这对于地理空间数据的可视分析也不例外. 随着移动采集技术和数据存储能力的不断提高, 地理空间数据所覆盖的空间范围越来越广, 时间跨度越来越大. 在分析如此大规模的地理空间数据时, 过滤常常是必不可少的操作. 地理空间数据可视化中的过滤是既普遍又特殊的: 一方面, 过滤作为一种数据操作, 将目标数据缩小定位到感兴趣的范围, 去除无关数据以加快计算速度、提高分析效率, 因此在可视分析中具有普遍性; 另一方面, 由于地理空间数据所特有的空间信息和由此而存在的 2D、3D 的空间关系, 地理空间数据可视化中的过滤设计又存在着特殊性. 因此, 本文对地理空间数据可视化中的过滤设计及相关工作进行了总结.

1 过滤及其讨论框架

过滤作为一种常用的数据操作, 在可视分析工作中得到了颇为广泛的应用. 然而, 不同的工作对过滤这个概念的定义并非完全一致. 作为调研的基础, 本文先确定过滤的定义, 从而划定调研的范围; 再基于此, 提炼出讨论相关文献的框架.

1.1 过滤

在可视化与可视分析领域, 研究人员对过滤给出了诸多的定义: Shneiderman^[10]提到过滤是排除不感兴趣的数据, 允许用户对显示内容进行控制, 从而更方便地关注于感兴趣的内容; Brehmer 等^[11]提到过滤是为可视化中元素进行包含和不包含条件的调节; Bartram 等^[12]提出过滤是排除不相关数据的过程; Andrienko 等^[7]说明过滤就是选出数据中相关的部分, 忽略不相关的部分; Roth^[13-14]从地理可视化的角度定义为鉴别出满足一定特征的地图特征. 这些定义都将过滤指向为一种筛选出满足一定条件的数据操作. 基于此, 本文确定过滤操作所包含的两部分: 过滤对象和过滤条件. 其中, 过滤对象即为过滤操作所作用的对象, 过滤条件则为过滤对象需要满足的约束条件. 过滤则为针对过滤对象设定过滤条件, 选出过滤对象中符合过滤条件的部分.

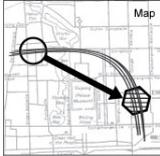
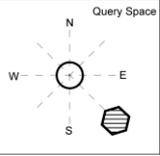
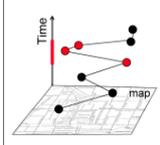
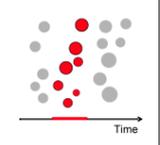
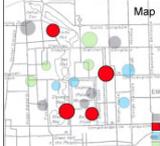
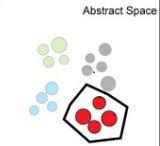
另外, 查询、检索与过滤息息相关. D'Ulizia 等^[15]

解释查询为句法层面上使用特定的语法定义从输入到输出的映射, 检索为在一个数据集中获得与某一信息相关的数据. 本文将查询、检索归并为过滤的一种, 即查询/检索是数据空间中满足一定查询/检索条件的过滤. 因此, 本文在进行调研时, 也将与查询、检索相关的工作考虑在内.

1.2 讨论框架

将过滤对象和过滤条件作为讨论过滤的 2 个维度, 如表 1 所示, 地理空间数据可视化中的过滤按以下 6 类进行讨论.

表 1 基于过滤对象和条件的讨论框架

过滤对象	过滤条件	
	基于地图空间	基于抽象空间
空间特征		
时间特征		
属性特征		

在地理空间数据的过滤设计中, 过滤对象按照地理空间数据包含的不同特征可分为空间特征、时间特征和属性特征. 地理空间数据是带有地理坐标的数据, 是对物体的空间特征、时间特征和属性特征的数字描述. 其中, 空间特征表现为物体的定位特征(物体的空间分布特征)以及物体之间的空间关系; 时间特征是物体的一种特殊的、自然的属性特征, 描述了物体随时间变化情况; 而将其他描述物体的抽象特征, 如数量、质量等统称为属性特征. 根据这 3 种不同的过滤对象, 本文分别讨论对应的过滤方法.

过滤条件的设定一方面取决于过滤对象的选择, 如对空间特征的过滤条件往往包含了定义空间关系的约束, 而针对属性特征的过滤条件可能只是数值范围的限定. 另一方面, 过滤条件的设定往往与可视化背景相结合, 以此更好地将过滤融入到整个可视分析的探索流程中. 对于地理空间数据可视化而言, 其可视化背景可以分为基于地理空间的可视化和基于抽象空间的可视化. 前者

利用地图等地理空间, 将空间数据按照空间位置进行布局以进行分析; 后者将数据编码到抽象空间中, 如散点图、高维投影等. 因此, 将基于地理空间和抽象空间作为过滤条件的 2 个子维度对地理空间数据的过滤进行讨论.

2 基于地图空间的过滤

在地理空间数据可视化中, 利用 2D, 2.5D 或 3D 的地理空间作为探索环境能直观地展现数据的空间信息. 为更好地与分析流程整合, 地理空间数据中的过滤常常内嵌于地理环境中. 将 3 种不同特征作为过滤对象, 其过滤的设计有所不同: 针对空间特征的过滤利用地理空间对空间进行直观的定位, 进而有效地设定空间过滤条件; 针对时间特征、属性的过滤则与时间或属性在地理空间中的编码方式紧密相关.

2.1 针对空间特征的过滤

通常, 基于地理空间的可视化将数据编码为点、线等图元, 并将其显示在以地图为代表的地理空间中. 空间特征的过滤条件所涉及的位置、区域大小、空间关系等, 可以直观地在地图空间中进行指定. 根据过滤条件的复杂程度, 本文从圈选过滤区域、定义过滤区域间关系和过滤滤镜 3 个方面来阐述针对地理空间中空间特征的过滤.

在地理空间中指定空间过滤条件最为直接的方式即为在地图上选择一定的空间区域, 过滤出与该区域满足一定空间关系的数据. 如图 1a^[16]所示, 通过矩形圈选过滤出经过特定码头的船舶轨迹数据. 在圈选的基础上, 还常常结合一定的空间语义, 如 Wang 等^[17]设计的基于道路的轨迹过滤, 通过圈选道路经过的区域, 过滤出图 1b 所示与道

路存在 4 种拓扑关系的轨迹. 此外, 圈选区域还与其他交互相结合的方式来完成逻辑更为复杂的过滤. 例如, Hurter 等^[18]在圈选区域之后, 支持用户通过“拖—放”的交互操作来实现集合逻辑运算. 图 1c 所示为飞机轨迹的可视化, 该方法支持圈选特定区域、再使用拖—放的交互来获得符合过滤条件的数据集和其补集.

在直接圈选过滤区域的基础上, 通过定义区域间的关系来设置多区域间的复合过滤条件. 例如, 通过定义区域间的先后顺序来过滤出满足一定方向的交通流. 如图 2a 所示, Ferreira 等^[19]设计支持任意形状的空间过滤窗口, 允许指定区域间的流向关系(图中带箭头线表示方向), 进而对出租车源—目的地数据进行过滤. 此外, 还支持将圈选组成过滤组, 进行组内取交、组间取或的集合运算. 如图 2b 所示, Lu 等^[20]设计的环形空间轨迹过滤允许用户通过直接操纵过滤器来修改过滤条件, 其中包括轨迹与区域之间的 6 种几何关系; 同时通过拖—放操作实现在过滤器间交通流方向的定义.

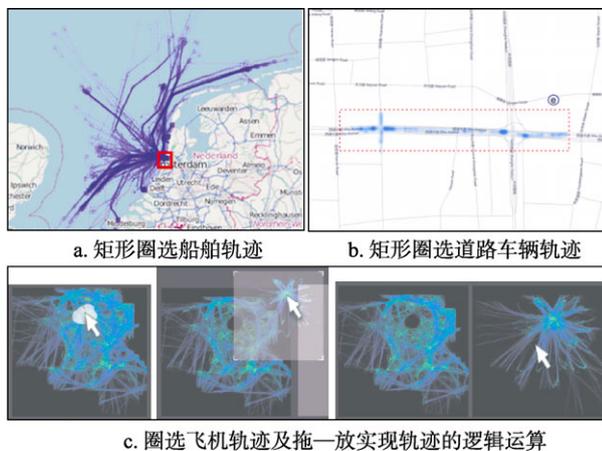
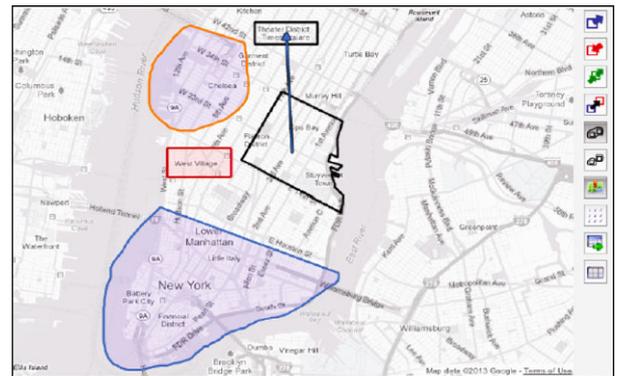
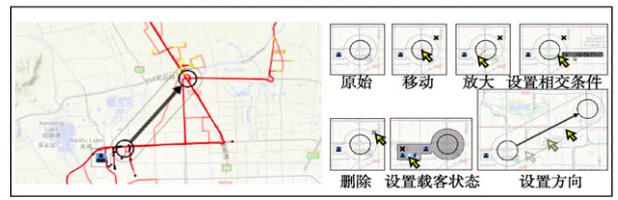


图 1 直接定义过滤区域



a. 任意形状区域间的交通流方向



b. 交通流与区域间关系的过滤

图 2 定义区域间关系的过滤

以上圈选区域作为过滤条件, 实现了离散区域的过滤查询. Turkey 等^[21]引入 3 种区域间的连续变化规则, 即随指定路径、大小、粒度连续变化实现了空间上的连续过滤. 如图 3 所示, 通过设置一定大小的空间窗口和一条路径, 过滤出位于沿该路线连续移动的空间窗序列中的数据.

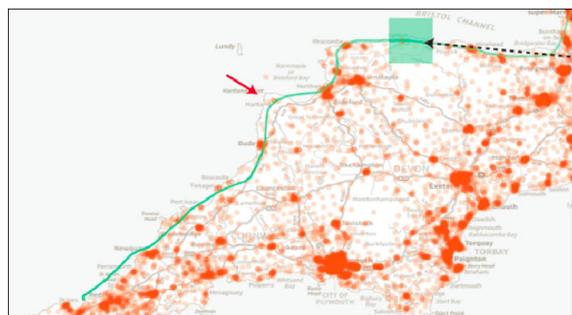


图 3 沿定义路径连续区域过滤

除直接圈选来指定过滤空间区域外, 基于地图的空间过滤常常使用滤镜(Lens)技术. 在可视化中, 滤镜技术^[22]借喻于现实中的镜片, 是一种支持选择子数据并将其进行数据变换后展示出来的技术. 当滤镜在 2D 空间中滑动时, 被滤镜所覆盖的区域会进行预先定义的数据转换, 并动态更新. 在地理空间中, 将数据变换特定为过滤操作, 滤镜自然地设计为在空间数据中实现过滤操作的载体^[22]. 早在 1994 年, Stone 等^[23]设计出广义的空间滤镜 MagicLens. 图 4a 是显示道路结构的滤镜. 随后, 研究人员在滤镜的实用性和交互性上做出了很多设计^[24]. 如图 4b 所示, 将滤镜在现实空间中用更为先进、自然的设备实现, 通过自然交互方式来对虚拟空间中的滤镜进行移动和摆放.

与 MagicLens 相比, Krüger 等^[25]设计的 Trajectory-Lenses 是针对轨迹数据的环形过滤器. 根据轨迹与环形过滤器不同的几何关系, 过滤器分为源、经过和目的 3 种. 如图 4c 所示, 在源或目的的滤镜中, 轨迹的起始点或终止点都会被显示出来, TrajectoryLenses 支持多组过滤器的集合操作. 不同组的过滤器用颜色区分.

2.2 针对时间特征的过滤

时间特征是地理空间数据中一种重要的属性

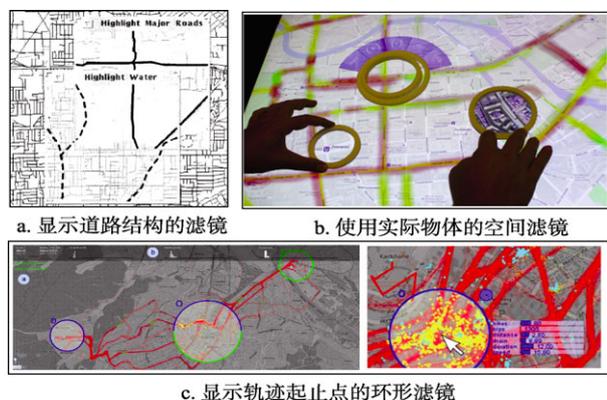


图 4 过滤滤镜

特征. 地理空间是 2D 或 3D 的, 而时间是 1D 的、线性的. 在地理空间中展示时间信息并不容易. 时空立方体^[26]是一种经典而直观的在地理空间中展现时间特征的可视化形式, 其中, 地图展现在 xy 平面内, 时间维度按 z 轴方向延伸, 从而在 2.5D 或 3D 的时空空间内可视化数据. 在时空立方体中, 针对时间特征的过滤通过在 z 轴上设计交互操作实现. 如图 5 所示, Tominski 等^[27]用条带表示其他属性的轨迹, 并将轨迹在 2.5D 的地图空间中堆叠起来, 通过在 z 轴上下滑动来定位时间过滤出对应时间的轨迹.

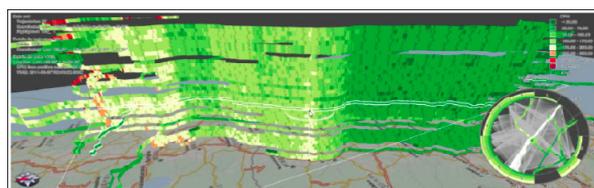


图 5 拓展 z 轴为时间的时空立方体

2.3 针对属性特征的过滤

在地理空间中, 属性特征通常按除空间位置之外的视觉元素编码, 如颜色、大小、方向、形状等. 因此, 针对属性特征的过滤操作, 通常是在属性—视觉映射上定义约束过滤条件来完成.

与空间特征相同, 用户可以通过直接选择在地理空间中一定视觉属性的数据来实现过滤. 例如, Andrienko 等^[28]用不同方向的柱形来表示米兰各区域交通流方向分布, 柱形的长度表示相对应方向流量的大小; Liu 等^[29]用点边图展现城市中区域间路径的多样性, 节点大小表示对应区域内的路径多样性. 以上 2 个应用支持用户点击特定大小的柱形或节点来选择与其相关的空间数据. 再如图 6 所示, Scheepens 等^[30]设计图标表示区域中船舶种类及对应的数量信息, 用户通过观察图标来确定感兴趣船舶种类或数量的区域, 点击图标来过滤出该区域中的船舶.

除通过直接过滤某一类特定属性外, 基于

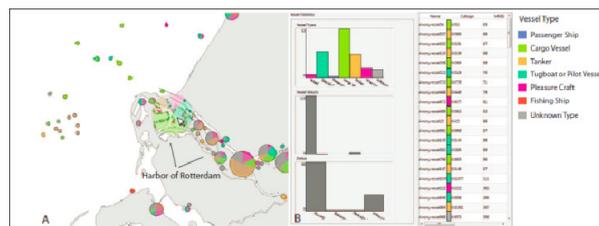


图 6 选择图标的过滤

视觉属性映射表的过滤可以获得多个属性特征范围的数据。颜色映射表是一种常见的属性视觉映射表，如图 7 所示，可以通过勾选代表船类别属性映射表中各项来过滤出不同类型的船只^[7]。类似地，Tominski 等^[27]通过选择代表不同数值范围的颜色映射来实现数值过滤。除了视觉编码和视觉映射表这样的显性属性特征外，还存在隐性的属性特征，如轨迹形状这样连续的几何属性特征。针对此类属性特征，Guo 等^[31]设计了 TripVista 在可视化路口交通轨迹时，使用描绘轨迹形状过滤出如图 8 所示相似形状的轨迹。

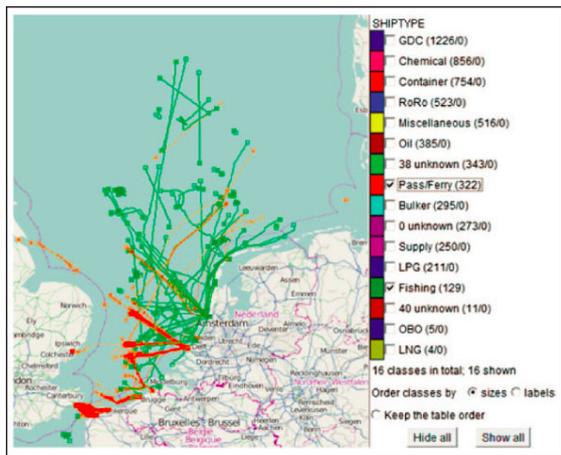


图 7 基于颜色映射表的过滤

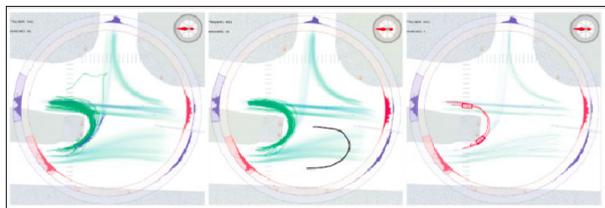


图 8 基于形状的过滤

3 基于抽象空间的过滤

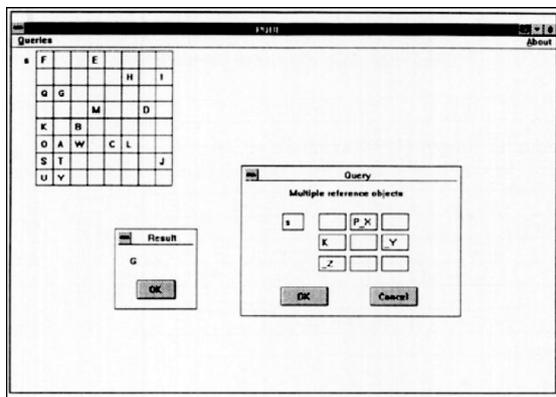
对于地理空间数据的可视化并非总是使用地理空间。将地理空间数据在抽象空间中进行可视化时，由于不受绝对地理位置的约束，空间特征的表达在一定程度上被弱化，而时间特征和属性特征则得到更为自由的表达。

3.1 针对空间特征的过滤

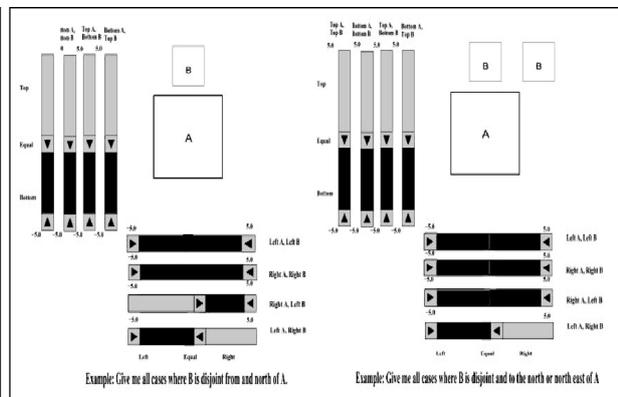
在抽象空间中，针对空间特征的过滤主要是通过空间视觉查询语言来完成。视觉查询语言按照一定句法规则定义从输入到输出的映射来查询数据。针对空间数据，空间视觉查询语言将空间元素之间的空间关系可视化，通过定义图元之间的位置关系、拓扑关系过滤出满足特定空间条件的数据。

Papadias 等^[32]设计了针对东、东西、西等 8 个方向关系的网格状查询语言。如图 9a 所示，通过在网格中摆放字母(下画线字母具有一定的过滤含义)来定义查询条件中物体间的位置关系。除了位置关系，Kaushik 等^[33]设计的视觉化查询还包含了空间物体的拓扑关系(如相交、相接等)，如图 9b 所示，通过代表不同量的滚动条，用户可以对空间关系定量地设定。

除以上通过控件(如滚动轴)设定图元间关系外，许多工作支持通过手绘这种更为自然的交互方式作为输入。Egenhofer 等^[34-35]将手绘输入(线、面等形式)转化为物体和物体之间的空间关系。再如图 10a 所示，用户通过手绘输入视觉化查询条件，再根据视觉化查询条件提取出一个五方位的场景网络，其表达了不同的信息：初步的相交关系、覆盖区域的各属性、数值、方位、量化的方位。图 10a

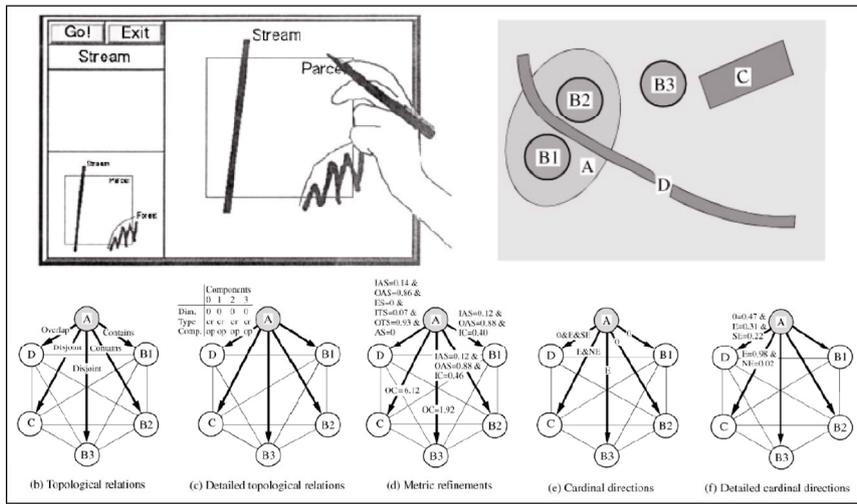


a. 基于网格定义空间位置关系

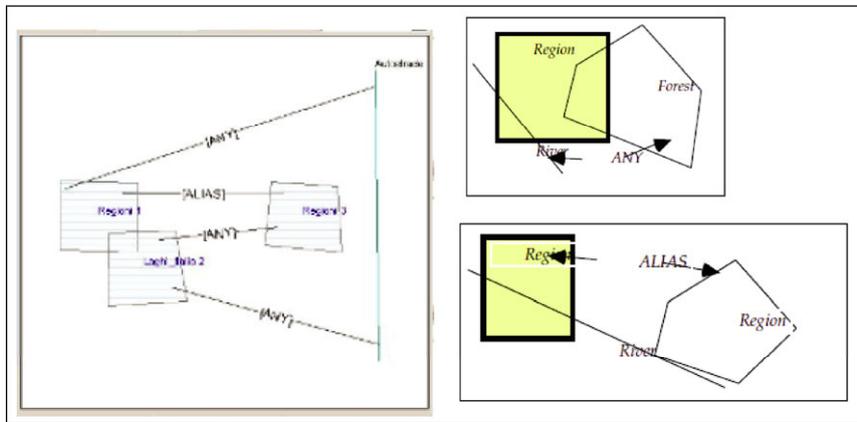


b. 基于滚动条定量设定空间位置

图 9 空间视觉查询语言



a. 定义包含相交关系在内的5种不同的场景网络



b. 定义额外2种操作(G-Any和G-Alias)解决二义性问题

图 10 手绘输入的视觉查询语言

下半部分显示的是这 5 个场景网络的子集, 即其他物体与物体 A 之间的关系。

在手绘作为输入时, 存在着二义性的问题需要考虑。如图 10b 所示, GeoPQL^[36] 中支持手绘输入; 同时针对在图元设置过程中存在的二义性问题, 它提出了 2 种额外的操作: 物体间的任意关系操作 (G-Any) 和物体别名操作 (G-Alias) (如图 10b 右部分所示)。类似地, Aufaure-Portier 等^[37] 设计了一套支持多种物体 (线、多边形) 的过滤操作 (集合操作、逻辑操作、结构操作及属性选择), 他们提出一种图模型, 将条件中所有可能存在的情况分层罗列在图的不同层面上; 用户在图中进行选择以解决二义性。此外, 他们通过维护一个用户的优先逻辑列表, 能自适应地给出最小二义性的情况。

3.2 针对时间特征的过滤

目前, 许多空间视觉查询语言已经扩展到时间特征。例如, D'Ulizia 等^[15] 定义了 3 种图元 (点、线、面), 12 种图元几何运算 (如并、交等), 7 种时间

运算 (包括覆盖、之前、相遇等) 以及 6 种空间运算 (如放大、缩小、消失等), 允许通过手绘输入的方式定义空间与时间的过滤条件。如图 11a 所示, 用户通过手绘输入了先后经过 4 个地点的时空查询条件, 其中每个平面表示不同的时间点。再如图 11b 中, 多边形面表示省, 输入的过滤条件是某个区域先前是 1 个省, 而后分为 2 个省。

另外, 时间轴是时间在抽象空间中的经典的可视化形式。在地理空间中针对时间特征的过滤常常在独立的时间轴上实现, 如图 11c 所示, Krüger 等^[25] 设计的 3 层时间规模的过滤器, 以及 Ferreira 等^[19] 设计的支持文本输入和时间选择的时间过滤器。此外, 时序视觉查询语言还支持设置事件之间的时序关系^[38-39]。如图 11d 所示, 不同的医疗记录可视化为不同的图标, 通过组合不同的图标、规定先后时序关系来过滤满足条件的病人。

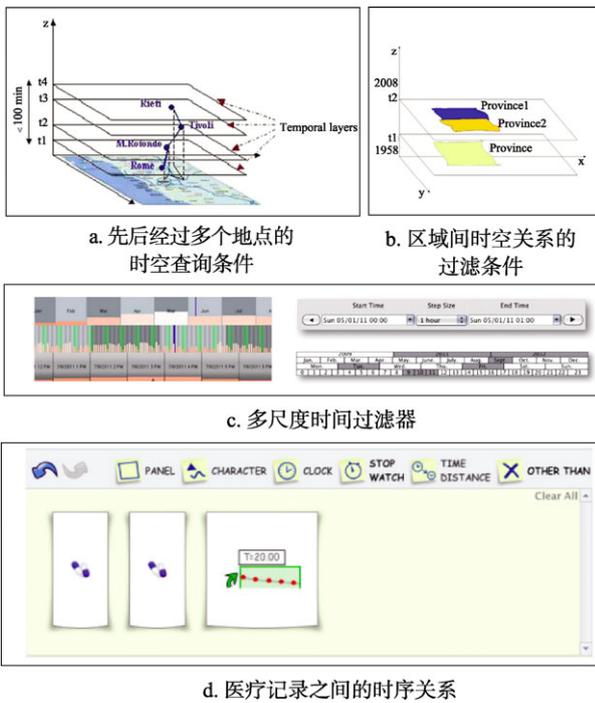


图 11 时间特征的过滤条件输入

3.3 针对属性特征的过滤

在抽象空间中，地理空间数据的属性特征可视化与其他数据的可视化相同。可能的可视化方法多种多样，因此过滤的设计也多种多样，本文不

一一赘述，仅列举出几个有代表性的地理空间可视分析中的属性特征过滤。

属性特征常常是随时间变化的，所以涉及到时间相关的可视化。如图 12a 所示，Guo 等^[31]基于主题河(ThemeRiver)^[40]的形式，用横轴表示时间，设计了展现一个路口的不同类型的移动物体的流量随时间变化。通过在主题河条带上嵌入表示方向的图标来展现车流方向的变化，用户可以通过点击特定车流方向的图标来过滤出对应的车流。Liu 等^[29]设计了一个环形可视化，其顺时针方向表示时间，沿时间方向将路径的速度等属性特征进行编码；用户可以刷选环形时间轴过滤出感兴趣的时间段，进而对相应的数据进行探索。

属性特征也可能是高维的，所以涉及高维数据的可视化形式，例如，平行坐标、高维投影等。Guo 等^[31]使用平行坐标(parallel coordinates)^[41]的形式来表现各属性上轨迹的分布，用户通过刷选轴上的范围来过滤出满足不同属性的轨迹。高维属性也可以采用降维的方式投影到低维度的抽象空间中，例如，Wang 等^[1]在出租车轨迹上提取出拥堵传播图，并根据图的相似度进行高维投影，用户通过圈选过滤出相似的拥堵传播图，如图 12b 所示。

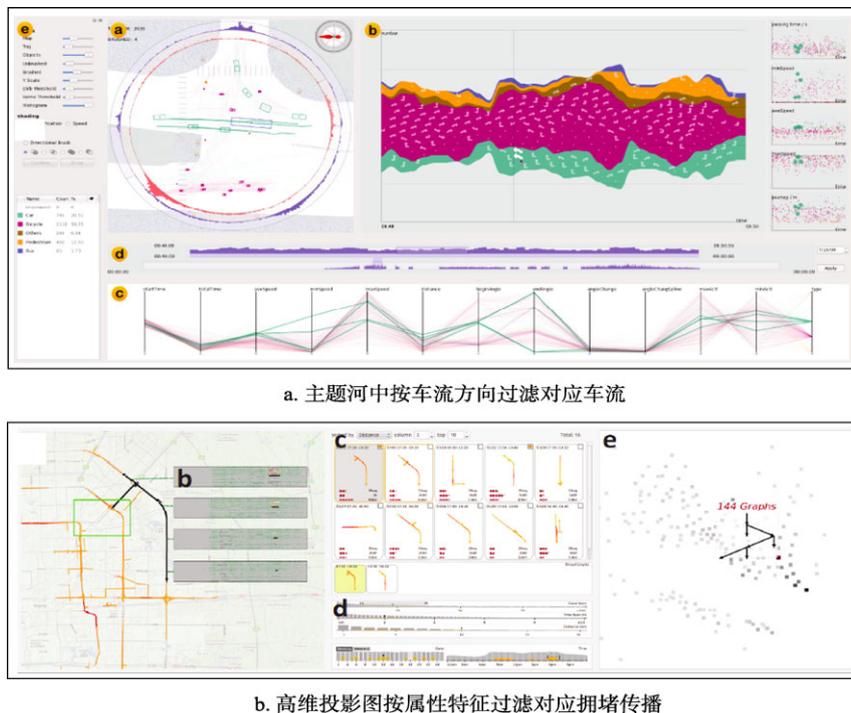


图 12 支持属性特征过滤的系统

4 讨 论

本文概述了在地理空间和抽象空间中针对地理空间数据的 3 种不同的特征的相关过滤工作. 下面对比这 2 种不同空间中的过滤设计, 并讨论对应的可视化适用场景. 另外, 单一特征的过滤通常不能满足复杂任务的需求, 因此还对过滤的复合进行讨论.

4.1 基于地理空间和抽象空间的过滤比较

基于地理空间和抽象空间的过滤设计各有优劣. 地理空间的过滤以地理空间作为探索环境, 直观地展示数据的空间信息. 在地理空间中, 用户能快速定位空间位置, 进而有助于进行空间特征的过滤. 同时, 在地理空间的可视化中, 时间和属性特征在空间中的分布也被直观地展现出来, 这将有效地支持用户设定它们与空间相关的过滤条件, 如过滤出在某一空间区域内的某类船舶^[30]等. 但也正因为空间关系的束缚, 在地理空间中对时间和属性特征的非空间相关过滤受到了一定的约束. 例如, 在地图上设定地点之间访问时间先后的过滤条件可能就会引起视觉遮挡问题. 相比之下, 由于不再受空间关系的束缚, 在抽象空间中设计时间特征和属性特征过滤的灵活度更大. 在抽象空间中, 空间特征的过滤多是用空间视觉查询语言表达. 空间视觉查询语言能很好地描述空间属性的相对空间关系, 但如果定义绝对的空间过滤条件, 往往需要辅助以额外的地理空间窗口.

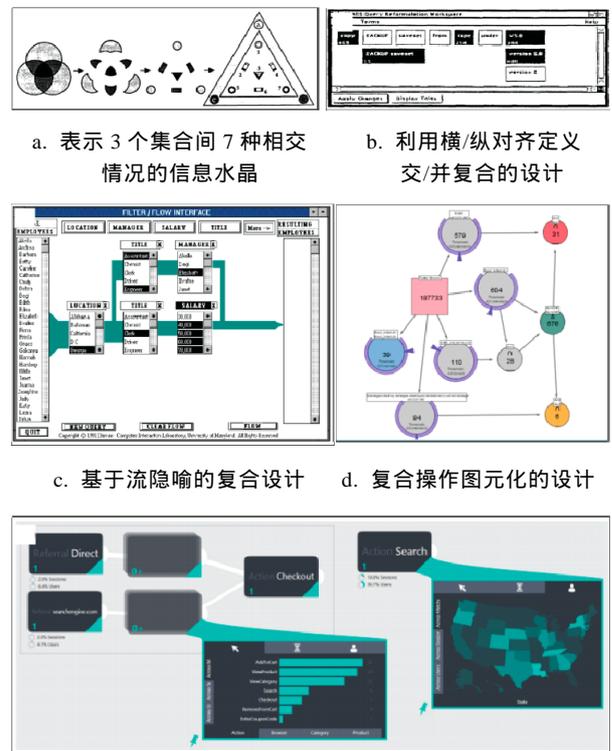
由于 2 种空间过滤具有的不同特点, 其适用的应用场景也不同. 基于地理空间的过滤紧密地耦合在地理空间中, 适用于以地理空间视口为主, 强调绝对地理位置的应用. 但是因为地理空间的约束, 将时间和属性特征的过滤融入到空间特征的过滤中, 设计难度高而且与应用比较相关. 而对于那些侧重于其他特征分析的地理空间数据应用, 将过滤设置在抽象空间中则有更大的灵活度. 以松耦合的方式将多种特征过滤复合起来, 相对地理空间的过滤而言, 它的设计难度更小、可拓展性较强, 但集成度较差.

4.2 过滤复合

随着分析任务复杂度的升高、分析数据种类的增加, 一个有效的可视分析系统常常既要有地理空间的可视化也有抽象空间的可视化. 因此需要考虑将二者的过滤复合起来.

根据现有工作, 本文将对过滤的复合以是否显性地定义复合操作为标准, 将复合分为显性和隐性 2 种.

显性的过滤复合将过滤图元化, 显性地对不同过滤之间如何结合进行定义. 将逻辑操作图元化有许多工作. 例如, 基于维恩图(Venn diagram)这种支持集合间布尔运算的所有可能的图形化表达, Michard^[42]设计了支持点击维恩图的不同部分而实现多集合间的布尔逻辑运算的功能. 但维恩图无法应对大量集合间关系的问题, 因此, Spoerri^[43]设计了信息水晶(InfoCrystal)的可视化查询形式, 将 n 个过滤条件所组成的 2^n-1 种交运算用图元来表示. 图 13a 为 3 个集合所有交运算的信息水晶产生过程: 3 个集合的 $2^3-1=7$ 种取交运算用多边形表示, 多边形的边数表示取交的集合数, 圆形表示 1 个集合, 用户可以直接通过选择这些图元来进行过滤. 信息水晶设计了一种复合运算——交, 还有工作考虑了交、并、差; 也有的方法是通过不同的布局来定义过滤条件之间的逻辑关系. 例如, Anick 等^[44]设计一个可视化操作平台, 将从自然查询语言中提取出的关键字及相近词可视化 2D 图元, 用户通过查看、添加以及对这些图元进行不同的布局来产生条件. 如图 13b 所示, 2D 图元的纵向排布表示并的关系, 横向排布表示交的关系. 再有, 将



a. 表示 3 个集合间 7 种相交情况的信息水晶

b. 利用横/纵对齐定义交/并复合的设计

c. 基于流隐喻的复合设计

d. 复合操作图元化的设计

e. 正则表达式的复合设计

图 13 显性定义的复合过滤

逻辑操作表现为图元, 例如, Young 等^[45]借鉴水流通过漏斗的比喻而设计了满足交(串联)、并(并联)及非(X 按钮)的视觉语言, 如图 13c 所示. 除了将逻辑关系的定义编码在组件的布局上这种方法外, Bosch 等^[46]是将这些逻辑关系定义在边上转移到了节点中, 即有特定的组件以组件的输出为输入, 其功能是对输入求交、求并或求差, 之后将得出的结果作为输出. 图 13d 中, 右侧有 3 个带有  标志的运算组件, 使得组件的布局更为自由. 如图 13e 所示, Zraggen 等^[47]设计了使用多个简单过滤器来描述、匹配一系列符合正则表达式句法的过滤条件.

隐性的过滤复合将多个过滤按照预先定义的操作进行运算. 交叉过滤机制(cross-filter)^[48]是目前广泛用于多窗口间的协同过滤的方案. 将多个窗口的过滤结果取集合交运算作为复合结果. 如 Wang 等^[1]的交通拥堵系统将多个属性特征的柱形图过滤, 并经过进行交运算后获得同时满足多个条件的拥堵事件.

5 结论与展望

近十几年来可视分析被逐渐应用于地理空间数据的分析中, 随着地理空间数据规模的不断增大, 过滤出相关数据、剔除无关数据对提高分析效率尤为关键. 本文调研了地理空间数据可视分析中过滤的相关工作, 并从过滤对象和过滤条件 2 个维度上进行讨论. 在基于地理空间的可视化中, 针对空间特征的过滤直观、有效; 相对应地, 在基于抽象空间的可视化中, 对非空间特征(即时间特征、属性特征)的过滤设计灵活、可拓展性高. 完成较为复杂的分析任务往往需要将空间特征和非空间特征的过滤结合在一起, 这是具有挑战性的工作. 目前的相关工作采取多窗口策略来支持任务分析, 并通过松耦合的方式将不同的过滤显性或隐性地复合起来, 这种方法能够做到有效; 但相比于在一个空间内进行统一的分析, 它并不是最直观的. 因此, 如何找到空间特征与非空间特征过滤的一个更好的契合点, 仍是一个值得探究的方向.

参考文献(References):

- [1] Wang Z C, Lu M, Yuan X R, *et al.* Visual traffic jam analysis based on trajectory data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2159-2168
- [2] Liu W, Zheng Y, Chawla S, *et al.* Discovering spatio-temporal causal interactions in traffic data streams[C] //Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2011: 1010-1018
- [3] Andrienko G, Andrienko N, Bosch H, *et al.* Thematic patterns in georeferenced tweets through space-time visual analytics[J]. *Computing in Science and Engineering*, 2013, 15(3): 72-82
- [4] Chen S, Yuan X R, Wang Z C, *et al.* Interactive visual movement patterns discovering from sparsely sampled geo-tagged social media data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2015, 22(1): 270-279
- [5] Ertl T, Chae J, Maciejewski R, *et al.* Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition[C] //Proceedings of the IEEE Conference on Visual Analytics Science and Technology. Washington D C: IEEE Computer Society Press, 2012: 143-152
- [6] Keim D, Andrienko G, Fekete J D, *et al.* Visual analytics: definition, process, and challenge[M] //Lecture Notes in Computer Science. Heidelberg: Springer, 2008, 4950: 154-175
- [7] Andrienko G, Andrienko N, Bak P, *et al.* Visual analytics of movement[M]. Heidelberg: Springer, 2013: 415
- [8] Pu Jiansu, Qu Huamin, Ni Lionel. Survey on visualization of trajectory data[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2012, 24(10): 1273-1282(in Chinese)
(蒲剑苏, 屈华民, 倪明选. 移动轨迹数据的可视化[J]. *计算机辅助设计与图形学学报*, 2012, 24(10): 1273-1282)
- [9] Wang Zuchao, Yuan Xiaoru. Visual analysis of trajectory data [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2015, 27(1): 9-25(in Chinese)
(王祖超, 袁晓如. 轨迹数据的可视分析研究[J]. *计算机辅助设计与图形学学报*, 2015, 27(1): 9-25)
- [10] Shneiderman B. The eyes have it: a task by data type taxonomy for information visualizations[C] //Proceedings of the IEEE Symposium on Visual Languages. Los Alamitos: IEEE Computer Society Press, 1996: 336-343
- [11] 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
- [12] Bartram L, Ware C. Filtering and brushing with motion[J]. *Information Visualization*, 2002, 1(1): 66-79
- [13] Roth R E. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2356-2365
- [14] Roth R E. Cartographic interaction primitives: framework and synthesis[J]. *The Cartographic Journal: The World of Mapping*, 2012, 49(4): 376-395
- [15] D'Ulizia A, Ferri F, Grifoni P. Moving GeoPQL: a pictorial language towards spatio-temporal queries[J]. *GeoInformatica*, 2012, 16(2): 357-389
- [16] Niels W, Huub V D W, Van W J J. Visualization of vessel movements[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(39): 15578-15583
- [17] Wang F, Chen W, Wu F A, *et al.* A visual reasoning approach for data-driven transport assessment on urban roads[C] //Proceedings of the IEEE Conference on Visual Analytics Science and Technology. Los Alamitos: IEEE Computer Society Press, 2014: 103-112
- [18] Hurter C, Tissoires B, Conversy S. Fromdady: spreading aircraft trajectories across views to support iterative queries[J]. *IEEE Transactions on Visualization and Computer Graphics*,

- 2009, 15(6): 1017-1024
- [19] Ferreira N, Poco J, Vo H T. Visual exploration of big spatio-temporal urban data: a study of New York city taxi trips[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2149-2158
- [20] Lu M, Wang A C, Yuan X R. Trajrank: exploring travel behaviour on a route by trajectory ranking[C] // *Proceedings of IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society Press, 2015: 311-318
- [21] Turck C, Slingsby A, Hauser H, *et al.* Attribute signatures: dynamic visual summaries for analyzing multivariate geographical data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2014, 20(12): 2033-2042
- [22] Tominski C, Gladisch S, Kister U, *et al.* A survey on interactive lenses in visualization[OL]. [2016-04-06]. https://www.researchgate.net/publication/266322178_A_Survey_on_Interactive_Lenses_in_Visualization
- [23] Stone M C, Fishkin K, Bier E A. The movable filter as a user interface tool[C] // *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 1994: 306-312
- [24] Ebert A, Weber C, Cernea D, *et al.* Tangiblerings: nestable circular tangibles[C] // *Proceedings of the Extended Abstracts on Human Factors in Computing Systems*. New York: ACM Press, 2013: 1617-1622
- [25] Krüger R, Thom D, Wörner M, *et al.* Trajectorylenses—a set-based filtering and exploration technique for long-term trajectory data[J]. *Computer Graphics Forum*, 2013, 32(3/4): 451-460
- [26] Kapler T, Wright W. Geotime information visualization[C] // *Proceedings of the IEEE Symposium on Information Visualization*. Los Alamitos: IEEE Computer Society Press, 2004: 25-32
- [27] Tominski C, Schumann H, Andrienko G, *et al.* Stacking-based visualization of trajectory attribute data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2565-2574
- [28] Andrienko G, Andrienko N. Spatio-temporal aggregation for visual analysis of movements[C] // *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. Los Alamitos: IEEE Computer Society Press, 2008: 51-58
- [29] Liu H, Gao Y, Lu L, *et al.* Visual analysis of route diversity[C] // *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. Los Alamitos: IEEE Computer Society Press, 2011: 171-180
- [30] Scheepens R, van de Wetering H, van Wijk J. Non-overlapping aggregated multivariate glyphs for moving objects[C] // *Proceedings of the IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society Press, 2014: 17-24
- [31] Guo H, Wang Z, Yu B, *et al.* TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection[C] // *Proceedings of the IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society Press, 2011: 163-170
- [32] Papadias D, Sellis T. A pictorial query-by-example language[J]. *Journal of Visual Languages and Computing*, 1995, 6(1): 53-72
- [33] Kaushik S, Rundensteiner E A. Sviquel: a spatial visual query and exploration language[C] // *Proceedings of the 9th International Conference on Database and Expert Systems Applications*. Heidelberg: Springer, 1998: 290-299
- [34] Blaser A D, Egenhofer M J. A visual tool for querying geographic databases[C] // *Proceedings of the Working Conference on Advanced Visual Interface*. New York: ACM Press, 2000: 211-216
- [35] Egenhofer M J. Query processing in spatial-query-by-sketch[J]. *Journal of Visual Languages & Computing*, 1997, 8(4): 403-424
- [36] Ferri F, Rafanelli M. GeoPQL: a geographical pictorial query language that resolves ambiguities in query interpretation[M] // *Lecture Notes in Computer Science*. Heidelberg: Springer, 2005, 3534: 50-80
- [37] Aufaure-Portier M A, Bonhomme C. A high level visual language for spatial data management[M] // *Lecture Notes in Computer Science*. Heidelberg: Springer, 1999, 1614: 325-332
- [38] Jin J, Szekely P. Interactive querying of temporal data using a comic strip metaphor[C] // *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. Los Alamitos: IEEE Computer Society Press, 2010: 163-170
- [39] Jin J, Szekely P. Querymarvel: a visual query language for temporal patterns using comic strips[C] // *Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing*. Los Alamitos: IEEE Computer Society Press, 2009: 207-214
- [40] Havre S, Hetzler E, Whitney P, *et al.* Themeriver: visualizing thematic changes in large document collections[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2002, 8(1): 9-20
- [41] Inselberg A. The plane with parallel coordinates[J]. *The Visual Computer*, 1985, 1(2): 69-91
- [42] Michard A. Graphical presentation of Boolean expressions in a database query language: design notes and an ergonomic evaluation[J]. *Behaviour & Information Technology*, 1982, 1(3): 279-288
- [43] Spoerri A. InfoCrystal: a visual tool for information retrieval & management[C] // *Proceedings of the 2nd International Conference on Information and Knowledge Management*. New York: ACM Press, 1993: 11-20
- [44] Anick P G, Brennan J D, Flynn R A, *et al.* A direct manipulation interface for Boolean information retrieval via natural language query[C] // *Proceedings of the 13th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Los Alamitos: IEEE Computer Society Press, 1989: 135-150
- [45] Young D, Shneiderman B. A graphical filter/flow representation of Boolean queries: a prototype implementation and evaluation[J]. *Journal of the American Society for Information Science*, 1993, 44(6): 327-339
- [46] Bosch H, Thom D, Heimerl F, *et al.* Scatterblogs2: real-time monitoring of microblog messages through user-guided filtering[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2022-2031
- [47] Zraggen E, Drucker S M, DeLine R. (s)queries: visual regular expressions for querying and exploring event sequences[C] // *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 2015: 2683-2692
- [48] Weaver C. Cross-filtered views for multidimensional visual analysis[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(2): 192-204