

基于Grasshopper平台的数据可视化在城市设计中的研究与实践

Research and Application of Data Visualization in Urban Design Based on Grasshopper

陈铭泽¹ 张洋^{2*} 杨玉冰³ 方智果⁴
CHEN Mingze¹ ZHANG Yang^{2*} YANG Yubing³ FANG Zhiguo⁴

(1.伦敦大学学院巴特莱特建筑学院, 伦敦 WC1E6BT; 2.上海交通大学设计学院, 上海 200240; 3.重庆道合园林景观规划设计有限公司, 重庆 401120; 4.上海理工大学艺术设计学院, 上海 200135)

(1. Bartlett School of Architecture, University College London, London, UK, WC1E6BT; 2. School of Design, Shanghai Jiao Tong University, Shanghai, China, 200240; 3. Chongqing DONEHOME Landscape Architecture Planning & Design Co., Ltd., Chongqing, China, 401120; 4. School of Art Design, University of Shanghai for science and technology, Shanghai, China, 200135)

文章编号: 1000-0283(2022)05-0044-08
DOI: 10.12193/j.laing.2022.05.0044.006
中图分类号: TU986
文献标志码: A
收稿日期: 2022-02-28
修回日期: 2022-04-04

摘要

随着数字时代的来临, 基于Grasshopper的数据可视化技术是近年来新兴的数据表达方法, 其作为城市设计者常用的基本可视化技术之一, 极大地提高了决策过程与设计成果的科学准确性、实操可行性及可持续适应性, 在城市设计中发挥着不可替代的作用, 并将持续作为城市设计的主导媒介。本研究概述了数据可视化发展历程及Grasshopper数据可视化在城市设计领域中的应用现状, 梳理了基于Grasshopper平台的数据提取和筛选的基本流程, 并对城市数据采集平台、数据格式和数据处理方法进行了比较。最后, 总结出“数据写入、数据整合和结构处理、可视化模型构建、可视化变量选择和处理”4步流程模式, 并应用于两个城市尺度的可视化实践——“伦敦各地区和英国主要地区家庭年平均收入数据可视化”和“伦敦主要树种数量及分布情况与对应家庭收入的可视化”。基于Grasshopper的数据可视化结果具有很强的表现力, 可用于场地背景挖掘、数据动态对比等方面, 为研究提供数据驱动基础与科学背景。Grasshopper进行数据可视化, 具有极强的可拓展性、可编程性、可修改性, 在数据驱动下的城市设计与规划领域具有良好的应用前景和开发潜力。虽然近些年Grasshopper的功能日渐强大, 但其在数据可视化领域仍有着不容忽视的缺陷, 仍需更多相关研究者和开发人员继续挖掘其深入潜力并进一步应用于实践。

关键词

数据可视化; 参数化设计; 可视化编程语言; 城市设计; 视觉分析

Abstract

With the advent of the digital era, Grasshopper-based data visualization technology has been an emerging data representation method in recent years. It greatly improves the scientific accuracy, practical feasibility, and sustainable adaptability of the decision-making process and design results. It plays an irreplaceable role in urban planning and design as one of the basic visualization techniques for urban planners. This study outlines the development history of data visualization and the current application status of Grasshopper data visualization in urban planning and design, compares the basic process of data extraction and screening based on the Grasshopper platform, and compares urban data collection platforms, data formats, and data processing methods. Finally, the four-step process model of “data preparation, data integration, and structural processing, visualization model construction, visualization variable selection, and processing” is summarized and applied to two urban-scale visualization practices: “visualization of average annual household income data in London and UK” and “number and distribution of major tree species in London with corresponding household income visualization.” The Grasshopper-based data

陈铭泽

1998年生/男/福建福州人/在读硕士研究生/助理研究员/研究方向为数据驱动景观与城市规划

张洋

1986年生/男/山东青岛人/博士/讲师/研究方向为智慧景观、城市生物多样性

杨玉冰

1997年生/女/四川乐山人/硕士/研究方向为以水文系统为基础的传统人居环境

*通信作者 (Author for correspondence)
E-mail: danielzhang86@sju.edu.cn

基金项目:

国家自然科学基金项目“基于开放数据的街道近人空间量化研究”(编号: 51808337); 教育部人文社会科学基金“基于‘产、文、景’一体化的乡村景观设计策略研究”(编号: 18YJC760012)

visualization results are highly expressive. They can be used for site context mining and dynamic data comparison, providing a data-driven foundation and scientific context for research. The extreme extensibility, programmability, and modifiability of data visualization using Grasshopper have good application prospects and development potential in the field of data-driven urban design and planning. Although Grasshopper has become more and more potent in recent years, it still has shortcomings in the field of data visualization that cannot be ignored. More researchers and developers are needed to further tap its potential and put it into practice.

Keywords

data visualization; parametric design; Grasshopper; urban design; visual analytics

1 数据可视化发展历程与应用现状

1.1 数据可视化的概念

数据可视化是指将大型数据集中并以图形图像的形式表示, 并利用数据分析和开发工具发现其中未知信息的处理过程^[1]。数据可视化最初是指对大型数据库或数据仓库^[2]中的数据可视化, 是可视化技术在非空间数据领域的应用, 使人们不再局限于通过关系数据表来观察和分析数据信息, 能以更直观的方式看到数据及其结构的深层次关系。数据可视化技术的底层逻辑是将一个数据库中的子数据项作为单个图元元素表示, 将大量的数据集可视化为直观的数据图像, 同时将数据的各个属性值以多维数据的形式表示, 可以从不同的维度观察数据, 从而对数据进行更深入的观察和分析^[2]。

1.2 数据可视化发展历程

大数据时代, 数据的分析过程往往离不开机器和人的相互协作与优势互补。人类从外界获得的信息约有80%以上来自于视觉系统^[3], 分析者面对大数据处理后的直观可视化图示时, 通常能够迅速洞悉数据背后的深层次信息, 并将其转化为知识和智慧^[4]。

普遍意义上的数据可视化被认为伴随统计学的诞生而出现, 从人们开始观察世界进而产生策略、管理等需求的时候, 用图形图像描绘、记录量化信息的思想就已经出现了。

从时间发展上看, 数据可视化的发展实质是伴随着数据本身意义的丰沛, 人们对于“数据”理解观念变化和可视化技术的发展。从空间发展上看, 数据可视化大类下又发展出相对独立而又紧密联系的新兴领域, 如科学可视化、信息可视化、知识可视化和可视化分析等(表1)^[4]。在大数据时代, 数据可视化技术不再仅仅是将散乱复杂、抽象多元的数据图像化; 也不再仅仅是提高技术效率与降低成本。实际是以“分析数据可视化的呈现结果、获取数据背后的隐含价值从而引导作出正确决策”的服务为导向, 而不断更新的数据基础决定了数据可视化也是一个持续的过程, 使用者能及时发现问题, 完善阶段性的决策^[5]。

1.3 数据可视化在城市设计中的应用现状

数据可视化在城市设计领域最常见的应用是将地图、平面布局、透视图等以往由徒

手绘制的图纸转变为数字化方式表达^[6]。城市设计中数字可视化的核心在于二维和三维形式的表示, 伴随着20世纪80年代初的计算机图形化技术的出现, 城市设计中的数字可视化首先应用于社区和建筑综合体的图示表达。目前, 中国有大量针对平方公里空间尺度的大尺度城市设计。如此的尺度规模并不是研究者能够用常规的城市设计概念、原则和技术方法可以掌控的^[7]。基于可视化技术的城市设计实践改进和扩展了以往城市设计专业的方法论和实践, 推动了城市设计的发展与城市建设管理科学化进程, 尤其是其数字化技术表示和语言能够与城市建设和管理技术平台有效结合。可视化技术在城市设计的方法论上具有革命性意义, 极大提高了设计成果和决策过程的准确性、科学性、可行性和适应性。与此同时, 可视化技术正在推动城市设计学科本身的发展, 为人们思考和理解空间方式带来巨大的转变。随着可视

表1 数据可视化发展历程表
Tab. 1 History of data visualization

时间 Time	发展情况 Developments
15世纪-17世纪	可视化思想起源, 数据可视化的早期探索阶段
18世纪	数据可视化初步发展, 直方图、柱状图、饼图、环图等开始出现
19世纪前半叶	数据开始得到重视, 数据图形在视觉表达方面有了极大进展
19世纪后半叶	数据可视化第一个黄金时期, 图形、图表被广泛应用
20世纪前期	黄金时期的数据表达方式已够用, 图形表达的研究并无新的进展
20世纪中后期至今	数据可视化依附计算机科学与技术拥有了新的生命力, 迎来第二个黄金时期

化技术系统和技术标准的不断完善,其呈现出集成化趋势。因此可视化现在并将继续成为城市设计的主要媒介。

目前一般的数据可视化工具只通过不同的显示方式对数据进行整合后提供给用户,用于观察分析数据与地理位置之间的关联性信息^[8]。然而,伴随着数据处理技术的不断进步,城市数据的表现已经不仅仅满足于基本的可视化工具对数据的图区以及简单的表达。为了适应城市数据量几何式增长所带来的高密度和高强度的分析需求,新的工具要能迅速实施获取、选择、分析、显示城市设计所需要的信息。因此,可视化工具必须满足易用性、实时性、多元数据集成融合等^[9]特性。

如今,Rhino与Grasshopper结合的参数化技术可极大程度满足城市设计研究的核心需求。Grasshopper是基于Rhino软件的插件,作为参数化设计中常用的工具之一,具有处理庞大数据的逻辑程序、节点模式可视化操作、实时效果呈现的特点,故基于Rhino与Grasshopper结合的数据可视化平台,对数据有着强大的处理管理和表达拓展能力。《Grasshopper参数模型构建》一书中讲述了Grasshopper工具的使用方法与大量案例索引^[10];在《Rhino&Grasshopper参数化建模》一书中详述了两种软件的软件特点和使用方式,并给出了二者结合建造模型的实例^[11];《参数化建筑设计》一书讲述了使用Grasshopper构建多种建筑类型的造型方法^[12];梁尚宇^[13]在滨水空间规划设计时首次利用Grasshopper参数化平台进行辅助;谭爵超^[14]在对景观桥和廊架等景

观构筑设计研究时也基于Grasshopper平台之上。城市设计师可以通过Grasshopper的逻辑构建系统来应对不同类型城市设计分析的需要^[15]。Grasshopper的底层逻辑是通过量化运算程序建立数据与几何物体之间的参数关系,利用几何容器、数据容器、几何运算器、数据运算器4类组件形成可调控的计算机模型。它在处理大量数据时,对于设计场地中可调控的与不可调控的因素通过系统的运作特征与关系,对数据进行转化与处理,将设计结果在Rhino平台中呈现。进而,对于实现城市设计的“第四代范型”——“从数据采集到数字设计,再从数字设计到数字管理”^[16],Rhino与Grasshopper结合的参数化技术平台将作为可持续的中坚力量。

2 基于Grasshopper的数据可视化的工作流程

2.1 数据提取和筛选

2.1.1 数据提取平台类别

数据提取是数据可视化的第一步,也是至关重要的过程之一。随着信息时代的到来,产生的数据量在成倍增长。量化数据可以通过不同平台获取,数据获取有开源和非开源两大类平台,对研究者而言是易获取的。

开源平台包括政府开源数据平台与主题数据开源平台。第一类平台数据往往是由国家政府或地方政府的有关部门收集公布,并且提供公共开放链接。例如英国开放数据平台^①、上海公共数据开放平台^②。这类数据涵盖类型往往具有宏观性、广泛性和全面性。常见分类包括商业和经济、政府、城镇

和城市、犯罪和司法、交通、健康、教育、地图、环境、社会等。第二类平台数据由政府相关部门、大学、专业协会、企业进行收集发布,大部分提供开放链接。例如英国林业开放数据^③由Forestry Commission运营,提供来自英格兰(England)、苏格兰(Scotland)、威尔士(Wales)的林业相关数据;中国地理空间数据云^④由中国科学院计算机网络信息中心运营,提供中国地区地理信息相关数据,包括数字高程(Dem)、陆地卫星遥感数据(Landsat)、光谱数据(Modis)等数据;开放街道地图(OSM)网站由英国非营利组织OpenStreetMap^⑤基金会赞助和维护,提供全球道路、土地使用、兴趣点、水路等数据。主题数据平台类数据拥有更强的专业性、精确性、针对性。

非开源平台可以通过编程语言或其他爬取工具进行网页提取的网站。Python作为大数据行业获取数据的核心技术,主要分为通用网络爬虫和聚焦网络爬虫两类^[17]。其中通用网络爬虫基本逻辑极其实实现原理如下:第一步,爬取网站初始URL,解析DNS得到主机IP地址;第二步,将爬取的数据存储到数据文件或数据库中,并获得新的URL;第三步,将新的URL存放到队列中,依据新的URL爬取网页与存储数据,并重复以上过程;第四步,满足停止条件时,停止爬取^[18]。而聚焦网络爬虫的抓取与执行流程比通用网络爬虫更加复杂,其可以在抓取的同时,对批量数据进行信息筛选、处理,快速过滤掉与主题无关的URL地址,并将与主题高相关性URL放入队列中,再多次循环重复抓取、筛选工作直

① <https://data.gov.uk/>

② <https://data.sh.gov.cn/index.html>

③ <https://data-forestry.opendata.arcgis.com/>

④ <http://www.gscloud.cn/>

⑤ OpenStreetMap provides map data for hundreds of websites, mobile apps and hardware devices (Archived from the original contenton, 2021-05-24)

至满足要求后再终止执行^[17]。另外Python也为网页数据爬取和数据分析提供了许多工具拓展包作为辅助，使数据爬取工作如虎添翼。如基于Python的BeautifulSoup可以快速高效地爬取网站数据，Pandas工具能方便灵活地清洗分析数据，调用Python的Matplotlib工具包能便捷地把数据分析结果图形可视化^[19]。

2.1.2 常见的数据格式和用途

常见数据形式可以分为两类。第一类直接文本类，这类数据往往可以直观地以文字形式呈现，包括：document、website、txt、pdf file、spread file、image、csv file（一种可以被excel读取的文件格式）。第二类间接转译类，这类数据常需要通过ArcGIS、Grasshopper等软件进行可视化表达，包括：Shapefile，一种可以在GIS系列软件读取或者Grasshopper进行读取的矢量文件格式，包括points、polylines、polygons，常用语储存城市POI、路网、用地类型等信息；Geo Tiff，一种可以在GIS系列软件和Grasshopper插件进行读取的栅格文件格式，常见的DEM数字高程数据、Landsat遥感卫星数据属于这种格式。

2.1.3 数据筛选和数据清理

通过上述平台获取的数据往往需要经过二次筛选和梳理才能通过可视化工具进行可视化表达。数据筛选和清理的目的是检测数据中存在的错误和不一致，以剔除或者改正它们，提高数据质量^[20]。数据处理的标准是简洁清晰、重点明确、操作性强，主要有两种方法：第一种处理空缺值，可以用回归、贝叶斯形式化方法工具或判定树归纳等确定空缺值；第二种平滑噪声数据（噪声数据是一个测量变量中的随机错误或偏差，包括错误的值或偏离期望的孤立点值），可以用“分

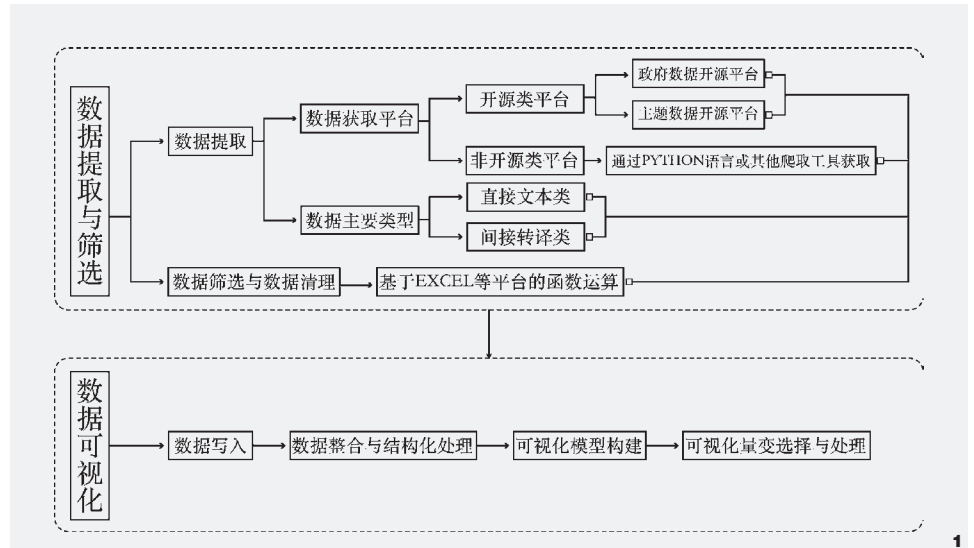


图1 数据可视化在景观与城市设计中的应用流程模型
Fig. 1 Process model for data visualization in landscape and urban design

箱”“回归”“计算机检查和人工检查结合”以及“聚类”等数据平滑技术来支持，识别和删除孤立点^[21]。

常见数据往往可以由Excel呈现，数据处理要根据自己的可视化目标进行科学化的处理，删除不需要的数据列，保留需要的数据列，通过计算增加原本没有但需要的数据列。通过分列和首行筛选、数据分列、删除重复等操作，可以快速保留需要的数据，保持表格的清晰直观。通过一些基本函数的操作，可以完成数据匹配、个数统计、条件统计、数学算法等基础数据处理。常见算法有VLOOKUP函数、Count函数、Count if函数、SUM函数等。VLOOKUP函数：当你需要在一个表格或一个范围内按行查找东西时，可以使用。例如，按区名查找家庭收入。Count函数：函数计算包含数字单元格的数量，并对参数列表中的数字进行计数。使用Count函数可以获得一个数字字段的条目数，该字段在一个数字范围或数组中。Count if函数：使用统计函数之一的Count if来计算符合某一标准的单元格

的数量。例如，计算出现在某一区的物种数量。SUM函数：SUM函数可以增加数值。可以添加单个值、单元格引用或范围，或三者的混合。在进行有效的数据过滤和清理工作后，就可以应用到可视化工具进行可视化处理和分析了。

2.2 基于Grasshopper的数据可视化应用

数据可视化是Grasshopper的重要应用领域之一，利用Grasshopper进行数据可视化具有很强的可拓展性、可编程性、可修改性。相对线上可视化网站或Excel等基本可视化工具，Grasshopper有更加丰富的可操作性组件，可以辅助研究者和设计师进行更加丰富多元的表达。相对python等基于代码的可视化平台，Grasshopper又具有结果呈现的直观性。Grasshopper进行数据可视化的流程分为：数据写入、数据整合与结构处理、可视化模型构建、可视化变量选择与处理4个步骤（图1）。

(1) 数据写入：利用Grasshopper基础组

Code	Borough	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	2009/10
E09000001	City of London	£65,120	£66,410	£69,800	£74,510	£76,200	£81,300	£83,050	£89,340	£89,680
E09000002	Barking and Dagenham	£22,930	£23,200	£24,240	£25,760	£26,260	£27,960	£28,510	£30,650	£30,750
E09000003	Barnet	£39,190	£39,470	£41,010	£43,320	£43,890	£46,430	£47,050	£50,240	£50,100
E09000004	Bexley	£30,060	£30,420	£31,760	£33,730	£34,370	£36,560	£37,260	£40,020	£40,130
E09000005	Brent	£29,430	£29,500	£30,520	£32,120	£32,430	£34,200	£34,560	£36,820	£36,620
E09000006	Bromley	£37,110	£37,640	£39,380	£41,880	£42,700	£45,450	£46,340	£49,770	£49,900
E09000007	Camden	£47,940	£48,430	£50,470	£53,450	£54,270	£57,520	£58,390	£62,450	£62,330
E09000008	Croydon	£31,740	£32,040	£33,380	£35,340	£35,890	£38,040	£38,630	£41,230	£41,280
E09000009	Ealing	£34,000	£34,120	£35,340	£37,210	£37,570	£39,620	£40,030	£42,610	£42,360
E09000010	Enfield	£30,750	£30,830	£31,900	£33,570	£33,890	£35,730	£36,100	£38,430	£38,210
E09000011	Greenwich	£28,290	£28,830	£30,310	£32,390	£33,200	£35,520	£36,390	£39,290	£39,580
E09000012	Hackney	£26,320	£26,900	£28,380	£30,430	£31,290	£33,580	£34,520	£37,380	£37,770
E09000013	Hammersmith and Fulham	£43,100	£43,670	£45,640	£48,470	£49,350	£52,440	£53,370	£57,220	£57,250
E09000014	Haringey	£31,510	£31,870	£33,250	£35,280	£35,890	£38,120	£38,790	£41,600	£41,640
E09000015	Harrow	£35,720	£35,900	£37,240	£39,270	£39,730	£41,970	£42,480	£45,330	£45,160
E09000016	Havering	£30,610	£30,910	£32,210	£34,140	£34,710	£36,860	£37,510	£40,230	£40,290
E09000017	Hillingdon	£32,140	£32,360	£33,630	£35,530	£36,010	£38,110	£38,640	£41,300	£41,200
E09000018	Hounslow	£33,000	£33,120	£34,300	£36,120	£36,490	£38,490	£38,890	£41,420	£41,190
E09000019	Islington	£37,560	£38,050	£39,750	£42,220	£43,000	£45,710	£46,540	£49,920	£49,970
E09000020	Kensington and Chelsea	£75,210	£76,790	£80,820	£85,390	£86,480	£94,550	£96,700	£104,170	£104,690
E09000021	Kingston upon Thames	£40,100	£40,500	£42,180	£44,660	£45,340	£48,050	£48,770	£52,180	£52,100
E09000022	Lambeth	£33,280	£33,700	£35,200	£37,380	£38,060	£40,450	£41,170	£44,160	£44,200
E09000023	Lewisham	£28,930	£29,350	£30,720	£32,690	£33,360	£35,540	£36,270	£38,990	£39,130
E09000024	Merton	£39,270	£39,770	£41,530	£44,090	£44,870	£47,670	£48,500	£51,990	£52,020
E09000025	Newham	£21,270	£21,700	£22,860	£24,490	£25,160	£26,980	£27,720	£30,000	£30,310
E09000026	Redbridge	£32,250	£32,480	£33,760	£35,690	£36,190	£38,320	£38,880	£41,580	£41,520
E09000027	Richmond upon Thames	£50,450	£51,380	£53,940	£57,530	£58,800	£62,720	£64,040	£68,890	£69,140

图2 原始数据
Fig. 2 Raw data

Code	Borough	Measure	Year	Income
E09000001	City of London	Mean	2012/13	99390
E09000002	Barking and Dagenham	Mean	2012/13	34890
E09000003	Barnet	Mean	2012/13	54530
E09000004	Bexley	Mean	2012/13	44430
E09000005	Brent	Mean	2012/13	39630
E09000006	Bromley	Mean	2012/13	55140
E09000007	Camden	Mean	2012/13	67990
E09000008	Croydon	Mean	2012/13	45120
E09000009	Ealing	Mean	2012/13	45690
E09000010	Enfield	Mean	2012/13	41250
E09000011	Greenwich	Mean	2012/13	44370
E09000012	Hackney	Mean	2012/13	42690
E09000013	Hammersmith and Fulham	Mean	2012/13	62910
E09000014	Haringey	Mean	2012/13	45860
E09000015	Harrow	Mean	2012/13	49060
E09000016	Havering	Mean	2012/13	44430
E09000017	Hillingdon	Mean	2012/13	44950
E09000018	Hounslow	Mean	2012/13	44490
E09000019	Islington	Mean	2012/13	54950
E09000020	Kingston upon Thames	Mean	2012/13	116350
E09000021	Lambeth	Mean	2012/13	56920
E09000022	Lewisham	Mean	2012/13	48610
E09000023	Merton	Mean	2012/13	43360
E09000024	Newham	Mean	2012/13	57160
E09000025	Redbridge	Mean	2012/13	34260
E09000026	Richmond upon Thames	Mean	2012/13	45380
E09000027	Southwark	Mean	2012/13	76610
E09000028	Sutton	Mean	2012/13	48000
E09000029	Westminster	Mean	2012/13	49170

图3 清洗处理后的数据
Fig. 3 Data after cleaning

件“path+read file”可以完成txt、csv格式数据的直接读取。利用Lunchbox等插件可以读取复杂的excel数据。(2) 数据整合与结构化处理：数据导入之后，可以利用“list”和“set”组件工具栏对数据进行系列处理，区分图例标注数据、控制可视化变量数据和其他数据，然后根据需求，对数据进行区间映射、数学运算、Graft/Flatten等操作。(3) 可视化模型构建：根据目标表达数据，选取并创造合适的模型构建，可基于但不限于柱状图(Bar Chart)、折线图(Line Chart)、饼状图(Pie Chart)、散点图(Scatter Chart)、气泡图(Bubble Chart)、雷达图(Rader Chart)、平行周线图(Parallel Axes)、树形图(Tree Diagram)等，利用Grasshopper中的几何运算器构建坐标轴、基础图形，进行可视化建模。(4) 可视化变量选择与处理：构建模型之后，结合处理后的数据进行外观的表达描绘，常见外观因子包括但不限于形状、大小、粗细、颜色、

密集程度等。在完成以上4个步骤之后，可以直接利用Rhino操作平台中View capture to file进行可视化图纸的导出。

3 城市尺度的应用实践实例

3.1 案例背景综述

伦敦，大不列颠及北爱尔兰联合王国首都，是欧洲第一大城市和最大的经济中心。大伦敦被分为伦敦市(City of London)及其周围32个自治市(Boroughs)。其中，卡姆登(Camden)、格林尼治(Greenwich)、威斯敏斯特(Westminster)等12个自治市为内伦敦区域。布伦特(Brent)、哈罗(Harrow)、贝克斯利(Bexley)等20个自治市为外伦敦区域。除伦敦外，家庭年平均收入数据还涵盖了英格兰地区(包括英格兰整体及其西南地区、西北地区、东南地区、东北地区)、苏格兰地区、威尔士地区、北爱尔兰地区以及英国全境的数据，下文统称为英国主要地区。本文的研

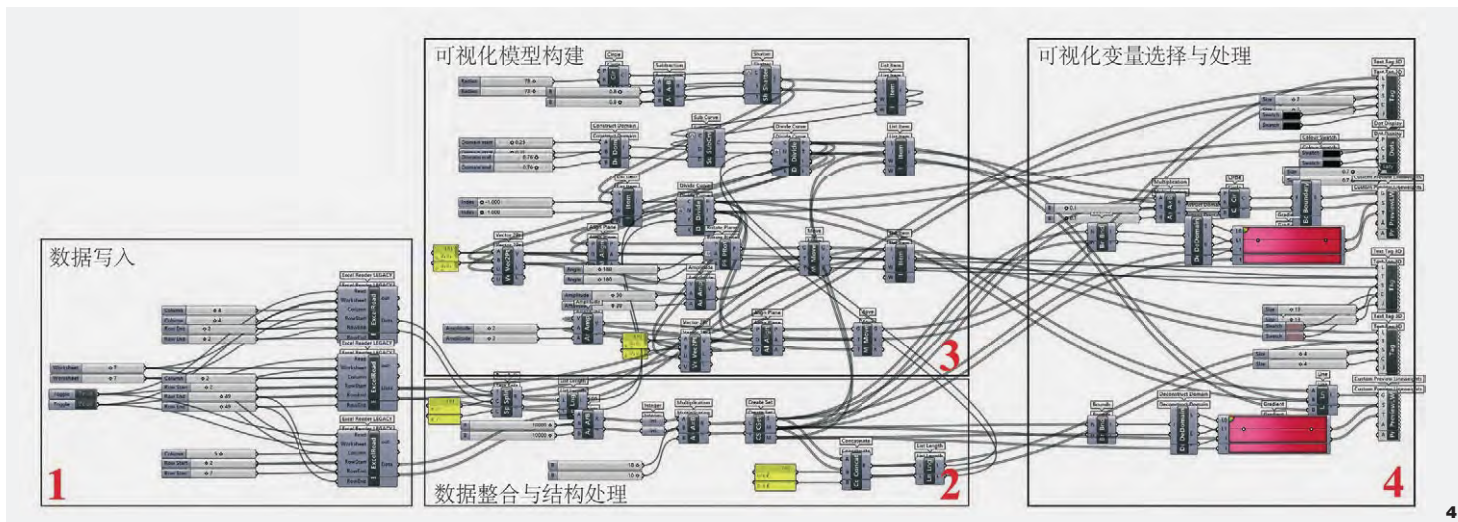
究案例基于该行政区划体系，以32个自治市为单位的伦敦各地区和英国主要地区进行数据收集、处理与研究。

3.2 伦敦各地区和英国主要地区家庭年平均收入数据可视化

实例一展示的是英国伦敦32个Boroughs和英国主要地区在家庭年平均收入的对比情况，其数据覆盖2001-2012年12年的时间。该数据来源于英国政府开源网站^①。下载的csv数据可以通过默认excel读取方式打开(图2)。通过简单的数据筛选和处理，删除与可视化无关的“Code”列表，并通过数据筛选功能，将2001-2012年中每年的数据进行筛选，以重组到相应的子表格中(图3)，这一步是利用“Lunchbox”插件批量读取多个表格的年份收入数据的基础。

通过观察筛选后的数据可以发现，收入可以通过10k(10 000英镑)一个区间进行整

① <https://data.london.gov.uk/dataset/household-income-estimates-small-areas>



合处理，收入与地区呈现多对一的关系，根据Grasshopper的既定逻辑：数据之间呈长映射关系进行连接，选择圆形、连线、颜色渐变作为可视化语言，按照数据写入、数据整合与结构处理、可视化模型构建、可视化变量选择与处理4个步骤，进行该案例的Grasshopper构建(图4)。

在数据写入步骤，通过“Lunchbox”——“Excel Reader”——“Number Slider”控制行列的选择并读取指定数据。在第二步“数据整合与结构处理”中，通过“Text Split”“List Length”等数据处理电池，将金额数据转化以千为计数的“XX K £”(XX千磅)格式。在可视化模型构建中，通过一系列的“几何运算器”组合，遵循Grasshopper中向量连线、映射区间等底层逻辑，完成根据半径画圆、根据对应关系连线等符号化语言的表达。最后在“可视化变量的选择与处理”中利用以“Gradient”电池为核心的组件完成颜色渐变等可视化操作。

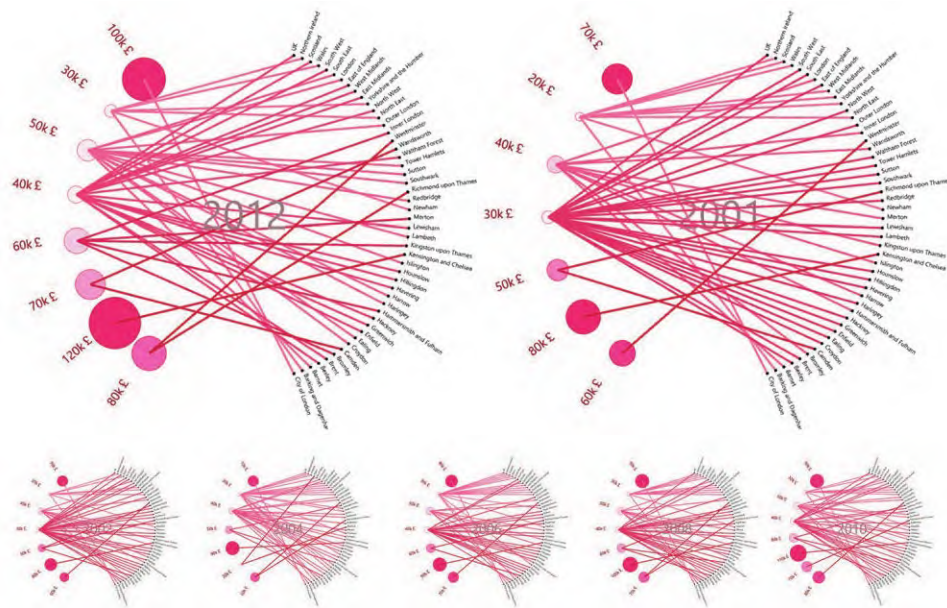


图4 Grasshopper四步骤电池图解
Fig. 4 Grasshopper's four-step battery illustration

图5 2001 - 2012伦敦各地区和英国主要地区家庭年平均收入数据可视化(部分展示)
Fig. 5 Visualization of average annual household income data for London and UK 2001-2012 (Partial display)

基于以上构建，只需调整输入图表数量的参数Slider，即可通过调整读取写入的数据，快速完成2001-2012年每年的数据可视化图纸(部分展示于图5)。

3.3 伦敦主要树种数量及分布情况与对应家庭收入的可视化

实例二基于实例一的家庭年平均收入数据，叠加伦敦市主要树种分布数据进行综合

① <https://data.london.gov.uk/dataset/local-authority-maintained-trees>

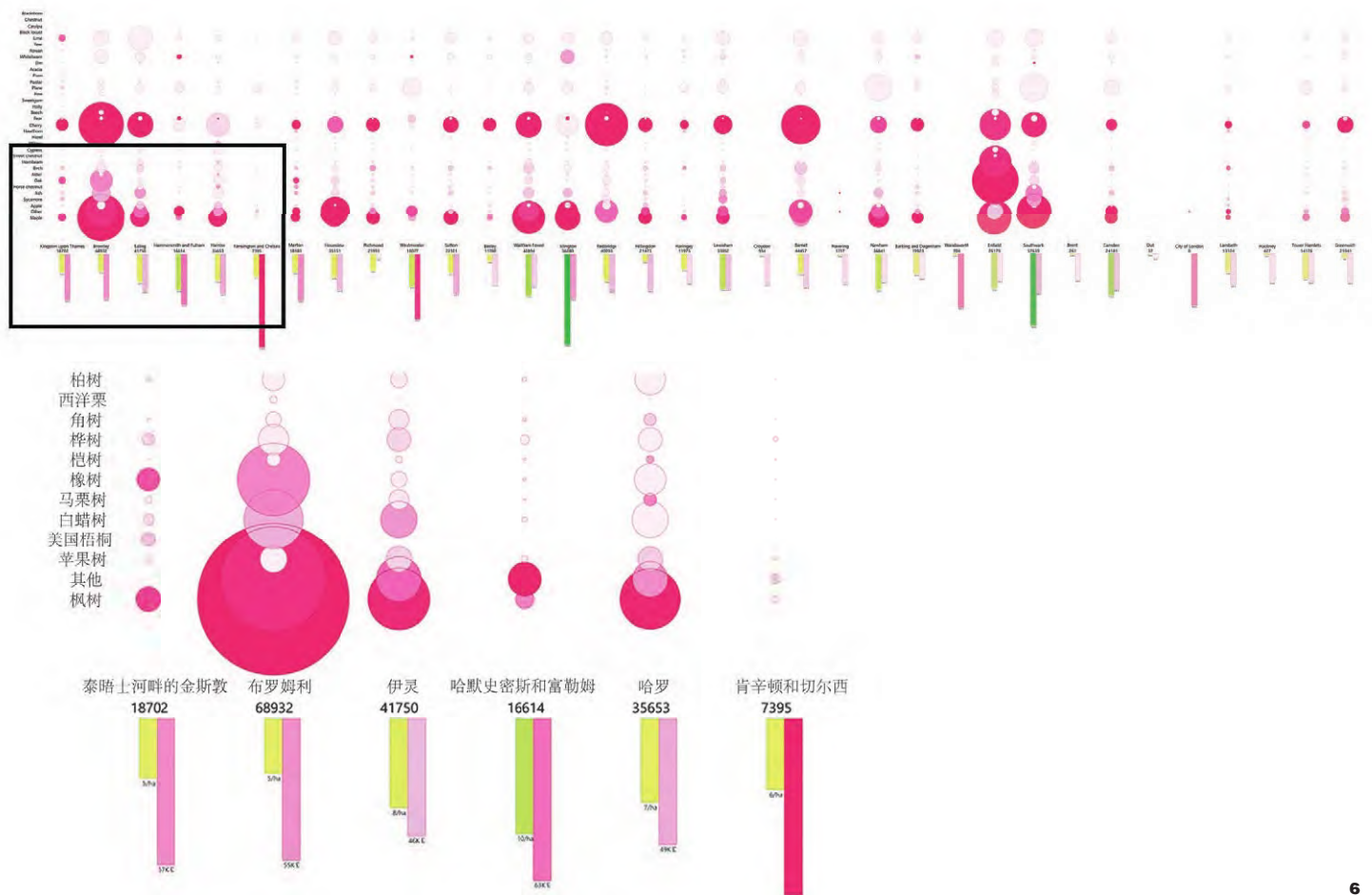


图6 伦敦主要树种数量及分布情况与对应家庭收入的可视化
Fig. 6 Visualization of the number and distribution of major tree species and corresponding household income in London

可视化展现，并在数据预处理中利用计算了各地区每 hm^2 拥有大树数量信息。数据来源于英国政府开源网站^①。可视化类型结合散点图和柱状图。在实际操作过程中，首先利用下载后的树种数据与实例一的收入数据进行结合处理，利用vlookup函数通过“Borough”属性列表进行数据匹配，然后通过countif函数对每个区的大树个数进行计数统计，利用基本运算计算各区每 hm^2 大树个数。同样通过Lunchbox插件读取导入数据，利用基本Grasshopper运算器构建X、Y坐标系：以Borough名称作为X轴方向变量，树种名称作为Y轴正

方向变量，收入和每 hm^2 大树数量作为Y轴负方向变量。通过圆形大小树种多少，柱状图中红色渐变与绿色渐变表达收入和每 hm^2 个数的高低，完成本实例的数据可视化及其局部放大结果(图6)。通过放大结果可以直观地看出Maple(枫树)是伦敦市数量最多的树木类型之一，kensington and Chelsea(肯辛顿和切尔西)地区拥有最高的家庭年平均收入，为伦敦市最富庶的地区。

3.4 可视化结果的导出与后期处理

完成可视化结果呈现后，可以基于Rhino

平台对图纸进行调整与导出。在Rhino6.0及更高的版本中，Grasshopper作为其内置的插件不需要进行额外的购买与安装。可以在Rhino界面中：“工具—选项—显示模式”进行背景颜色、物件属性、可见性等属性的调整。随后将可视化结果全部展示于显示视图中(常用Top视图)，再利用View Capture to File命令进行导出，并在该命令菜单对图纸要素进行设置。如有需要，也可以通过导出选取的物件将可视化成果转化为矢量格式，于Adobe Illustrator等矢量处理平台进行编辑，或进入Adobe Photoshop等图片处理平台对所需

的表达进行调整。

基于Grasshopper的数据可视化结果具有很强的表现力,可用于场地背景挖掘、数据动态对比等方面,为研究提供科学的数据驱动基础。上述两个实践案例生动直观地展现出伦敦以自治市(Boroughs)为单位的家庭年平均收入、主要树种分布等人文与自然因素信息。在时间上,可以横向动态对比10年内的数据变化情况;在空间上,可以直观看出不同区域的数据差异。这为后续研究根据规划设计目的进行场地选址、适应性分析、相关性分析等提供了良好的数据基础。

4 结论和讨论

利用Grasshopper进行数据可视化具有很强的可拓展性、可编程性、可修改性,在数据驱动下的城市设计与规划领域具有良好的应用前景和开发潜力。研究根据城市数据的分类入手,将城市数据分为政府开源数据平台与主题数据开源平台两类(非开源类数据及其获取方式不在本文研究范围内),并介绍其子分类与应用领域,继而概述了直接文本类与间接转译类两类数据类型,举例介绍其主要格式与分类依据。在数据筛选与清洗阶段,介绍了基于Excel函数的常见数据筛选方式,包括数据匹配、个数统计、条件统计、数学算法等基础数据处理。最后总结基于Grasshopper进行数据可视化的4步流程:数据写入、数据整合与结构处理、可视化模型构建、可视化变量选择与处理,并从数据获取、数据筛选与清洗、数据可视化表达,总结了流程模型可应用于其他多项实践领域。

通过对“伦敦各地区和英国主要地区家庭年平均收入数据可视化”和“伦敦主要树种数量及分布情况与对应家庭收入的可视化”两个实例的研究可以得出:Grasshopper在

效果表达具有很强的可塑性和可开发性,使用者即开发者,这是Grasshopper能够成为数据与图纸表现之间桥梁的关键原因之一。可视化技术在城市设计中正在发挥着对于方法论而言具有革新意义的重要作用,大大提高了设计成果及决策过程的准确性、科学性、可行性及适应性,并且将持续是城市设计的主导媒介。

虽然近年来开发者逐渐增多,随之而来的各种插件极速迭代更新,使得Grasshopper的功能也日趋强大,但其在数据可视化领域仍有着不容忽视的缺陷,如对数据的精确标注不够完全,科学统计分析没有R语言、Python、SPSS等软件实用性强等,仍需更多相关研究者和开发人员继续深入挖掘其在可视化领域的潜力并应用于实践。

注:文中图表均由作者绘制。

参考文献

- [1] 刘勘,周晓峥,周洞汝.数据可视化的研究与发展[J].计算机工程,2002(08):1-2.
- [2] INMON W H. Building the Data Warehouse[M]. New Jersey: Wiley, 2005
- [3] CARD S K, MACKINLAY J D, SHNEIDERMAN B. Readings in Information Visualization: Using Vision to Think[M]. San Francisco: Morgan-Kaufmann Publishers, 1999.
- [4] 任磊,杜一,马帅,等.大数据可视分析综述[J].软件学报,2014,25(09):1909-1936.
- [5] 曾悠.大数据时代背景下的数据可视化概念研究[D].杭州:浙江大学,2014.
- [6] 迈克尔·巴蒂,安德鲁·哈德逊·史密斯,陈宇琳.城市设计中的可视化分析、智慧城市与大数据[J].城市设计,2016(03):6-15.
- [7] 王建国,蔡凯臻.数字技术方法在现代城市设计中的应用[J].南方建筑,2008(02):28-32.
- [8] 蒋宏扬.基于大数据技术的城市设计方法创新[D].重庆:重庆大学,2016.
- [9] ANDERSON C. The Petabyte Age: Because More isn't Just More—more is Different[J]. Wired Magazine, 2008(07): 106-120

- [10] 包瑞清.参数模型构建[M].南京:江苏凤凰科学技术出版社,2015.
- [11] 曾旭东,王大川,陈辉. Rhinoceros&Grasshopper参数化建模[M].武汉:华中科技大学出版社,2011.
- [12] 高云河,白云生.参数化建筑设计[M].武汉:华中科技大学出版社,2016.
- [13] 梁尚宇.形式寻优:参数化辅助滨水空间形态规划设计[J].广东园林,2017,39(02):37-44.
- [14] 谭爵超.基于Grasshopper参数平台下的园林景观设计研究[D].广州:仲恺农业工程学院,2017.
- [15] 陈凌锋.基于Rhino与Grasshopper参数化技术在风景园林规划设计中地形的应用研究[D].广州:仲恺农业工程学院,2018.
- [16] 王建国.基于人机互动的数字化城市设计——城市设计第四代范型刍议[J].国际城市规划,2018,33(01):1-6.
- [17] 李玉香,王孟玉,涂宇晰.基于python的网络爬虫技术研究[J].信息技术与信息化,2019(12):143-145.
- [18] 黄兴荣,徐兴彬.基于Python爬虫技术和LDA模型的短文本获取技术分析[J].电大理工,2019(03):1-3.
- [19] 欧阳元东.基于Python的网站数据爬取与分析的技术实现策略[J].电脑知识与技术,2020,16(13):262-263.
- [20] 叶焕倬,吴迪.相似重复记录清理方法研究综述[J].现代图书情报技术,2010(09):56-66.
- [21] 刘莉,徐玉生,马志新.数据挖掘中数据预处理技术综述[J].甘肃科学学报,2003(01):117-119.