

武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: DP 算法自动实现方法—以河流化简为例
作者: 闫浩文, 张鑫港, 禄小敏, 李蓬勃
DOI: 10.13203/j.whugis20210412
收稿日期: 2022-07-13
网络首发日期: 2022-08-12
引用格式: 闫浩文, 张鑫港, 禄小敏, 李蓬勃. DP 算法自动实现方法—以河流化简为例 [J/OL]. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20210412>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20210412

引用格式：

闫浩文, 张鑫港, 禄小敏, 等. DP算法自动实现方法—以河流化简为例[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210412 (YAN Haowen, ZHANG Xingang, LU Xiaomin, et al. Approach to Automating the DP Algorithm—Taking River Simplification as an Example[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210412)

DP 算法自动实现方法—以河流化简为例

闫浩文^{1, 2, 3}, 张鑫港¹, 禄小敏^{1, 2, 3}, 李蓬勃¹

1. 兰州交通大学测绘与地理信息学院, 甘肃 兰州, 730070

2. 地理国情监测技术应用国家地方联合工程研究中心, 甘肃 兰州, 730070

3. 甘肃省地理国情监测工程实验室, 甘肃 兰州, 730070

摘 要: 曲线化简是自动地图综合的重要内容, 但其中广泛应用的 Douglas-Peucker 算法 (简称 DP 算法) 却是非自动化的, 原因是需要在算法执行之初由人工输入距离阈值 ε 。为此, 本文首先提出了一个多尺度曲线相似度的计算公式, 然后基于该公式, 以河流数据为例给出了地图比例尺与曲线目标相似度的函数关系推导方法和曲线目标相似度与 ε 的函数关系推导方法, 进而得出了 ε 与比例尺的函数关系, 由此实现了 DP 算法的自动化。实验研究表明, 利用本文提出的自动化 DP 算法可以获得指定地理区域不同比例尺的水系要素的化简结果, 化简结果与经验丰富的制图员的手工化简结果的相似度平均值为 0.927, 相似度总体表现良好, 表明了该方法良好的可靠性和较高的智能化。

关键词: 曲线化简; 地图综合; Douglas-Peucker 算法; 相似度; 河流数据

中图分类号: P228

文献标志码: A

收稿日期: 2022-07-13

第一作者: 闫浩文, 博士, 教授, 从事地图综合与空间关系研究。yanhw@mail.lzjtu.cn。

项目资助: 国家自然科学基金 (41930101)。

地图上 80%以上的要素（等高线、境界线、河流、道路和地类界等）以曲线形式表达^[1]。当地图由大比例尺变为较小比例尺时（此即地图综合），需要对曲线要素进行化简，其目的之一是为了满足地图上线状目标可视化的要求^[2-6]。曲线要素化简要求在尽量保持化简后曲线与原始曲线形状相似的情况下，使化简后曲线的复杂程度与目标地图的比例尺相适应^[7-10]。为了实现地图上曲线的化简，学者们提出了许多算法^[11-13]。其中，在地图综合中应用比较广泛的是 Douglas-Peucker 算法^[12]，本文简称其为 DP 算法。

在地图综合中，理想的曲线化简算法应该是自动化的，即其不受人工干预^[13]。但是，DP 算法并非自动，因为其在曲线化简开始时需要人工输入一个距离阈值 ε ($\varepsilon > 0$)。虽然已有改进的 DP 算法，例如适用于闭合曲线化简的 DP 算法、处理化简后曲线自相交问题的 DP 算法^[12]等，但它们均偏重于算法应用的普适性，并未考虑 DP 算法的自动化问题^[13-16]。

基于上述原因，本文将专注于地图上曲线状要素的化简问题，目标是提出一种 DP 算法的自动化实现方法。

1 DP 算法自动化的思路

1.1 问题描述

为了解决 DP 算法在曲线化简前需要人工输入 ε 的问题，有 2 种方式：（1）在算法执行前获得 ε 的值，把它嵌入到算法中；（2）找到 ε 的计算公式，把它编入算法。

对设想（1）而言，在面对特定地图综合任务时，根据已知的原始地图和目标地图比例尺，依靠有经验的地图专家，能够得到 ε 的值。但是，在地图数据实际生产中，需要顾及数量众多的地图比例尺和地理特点多样的制图区域，很难对 ε 进行枚举。

对于设想（2）而言，如果能找到 ε 的计算公式，把其融入 DP 算法中，就能免去手工输入 ε 的操作，实现算法的自动化。本文采用此设想。

2 基本思路

设有一个地图综合任务，原始地图和目标地图的比例尺分别为 S_0 、 S_1 ，原始地图上有 n 条曲线要素，欲借助 DP 算法化简表达达到目标地图上。根据前文讨论，要实现曲线的自动化简，就需要找到 ε 的计算公式。

在地图综合中， ε 的取值与地图比例尺和制图区域的地理特点有关。为了简化问题，对

于给定的地图综合任务而言，不考虑制图区域的地理特点（因为可视其为已知）。这样就有：

$$\varepsilon = f_1(S) \quad (1)$$

$$S = S_0/S_1 \quad (2)$$

其中， S 称为地图比例尺变化。

可用如下思路求 f_1 ：首先，地图综合是一种相似变换，故 DP 算法对曲线的化简可视为曲线在不同比例尺地图上的相似变换。其次，化简后曲线与原始曲线的相似度（用 Sim 表示）和地图比例尺变化 S 之间存在明显的单调函数依赖关系： S 越大，则 Sim 越小^[17,18]。所以，如果求得如下关系，就可以推导出 f_1 ：

$$Sim = f_2(S) \quad (3)$$

$$Sim = f_3(\varepsilon) \quad (4)$$

因此，下文首先探讨多尺度曲线相似度的计算方法，然后结合具体实验数据阐述比例尺变化与相似度的函数关系构建方法，接着讨论相似度与 ε 的函数关系，最后推导 ε 与比例尺变化的计算公式。

3 多尺度曲线相似度计算方法

对于给定的地图综合任务而言， S_0 和 S_1 已知，故 S 已知；但化简后不同比例尺地图上的曲线与原始地图上曲线的相似度 Sim 未知。所以，如何计算一条曲线在两个不同比例尺地图上的图形相似度和如何构建 f_2 是问题的核心。

计算两条曲线的相似度可以用欧氏距离、Manhattan 距离、Chebyshev 距离、Hausdorff 距离、Fréchet 距离、DTW(Dynamic Time Wrapping)等方法^[15-20]。顾及地图综合中曲线化简前后首末点保持不变的特点，前期研究发现，Hausdorff 距离更适合于多尺度地图上的曲线相似度计算^[21-27]。

设地图上线要素 A 、 B 的构成点集为： $A = \{a_1, a_2, \dots, a_m\}$ ， $B = \{b_1, b_2, \dots, b_n\}$ ，则其 Hausdorff 距离 $H(A, B)$ 定义为^[25,26]：

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (5)$$

其中，

$$h(A, B) = \max(a \in A) \min(b \in B) \|a - b\| \quad (6)$$

$$h(B, A) = \max(b \in B) \min(a \in A) \|b - a\| \quad (7)$$

其中， $\|\cdot\|$ 表示两点的距离范式，可以用欧氏距离、曼哈顿距离、切比雪夫距离等任何一种距离。本文选用欧氏距离。

设比例尺 S_0 的地图上有 n 个线要素集 $F_{S_0} = \{F_{S_0}^1, F_{S_0}^2, \dots, F_{S_0}^n\}$ ，化简到比例尺为 S_1 的地图上得到的对应要素集 $F_{S_1} = \{F_{S_1}^1, F_{S_1}^2, \dots, F_{S_1}^n\}$ ，则第 i 个要素与其化简后得到的对应要素的相似度为：

$$Sim(F_{S_0}^i, F_{S_1}^i) = 1 - \frac{H(F_{S_0}^i, F_{S_1}^i)}{MaxH(F_{S_0}^i, F_{S_1}^i)} \quad (8)$$

其中， $MaxH$ 是 $F_{S_0}^i$ 与 $F_{S_1}^i$ 构成的度量空间内的最大 Hausdorff 距离，计算方法如下：由于 DP 算法化简不改变曲线首末点，故可将 Hausdorff 距离的度量空间限定到原始曲线 $F_{S_0}^i$ 与化简曲线 $F_{S_1}^i$ 的最小面积外接矩形 $MABR$ (Minimum Area Boundary Rectangle)。由此容易得到 $MaxH(F_{S_0}^i, F_{S_1}^i)$ 为 $MABR$ 的两条边中的较长者（证明从略）。如图 1， ABC 为原始曲线， AC 为化简后曲线， ABC 的 $MABR$ 如虚点线所示，则 $MaxH(ABC, AC) = M_2$ 。

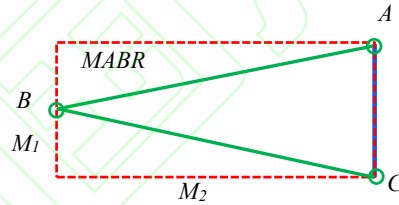


图 1 最大 Hausdorff 距离计算

Fig.1 Calculation of the maximum Hausdorff distance

n 条线要素在化简前后的整体相似性计算公式为：

$$Sim(F_{S_0}, F_{S_1}) = \sum_{i=1}^n \frac{Sim(F_{S_0}^i, F_{S_1}^i) \times L(F_{S_0}^i)}{L(F_{S_0})} \quad (9)$$

其中， $L(F_{S_0}^i)$ 是 $F_{S_0}^i$ 的长度， $L(F_{S_0})$ 是原始 n 条线要素的总长。

4 实现 DP 算法自动化的方法

4.1 S 与 Sim 的函数关系推导方法

为了得到比例尺变化和相似度之间的函数关系，选取某地区 1 幅 1:10K 矢量地形图上的 15 条线状河流作为实验数据（图 2）。由经验丰富的制图员对河流进行化简，得到了它

们在 1: 25K、1: 50K、1: 100K、1: 250K、1: 500K、1: 1M 共 7 种比例尺地图上的图形表达。此处，1K=1 000，1M=1000 000。

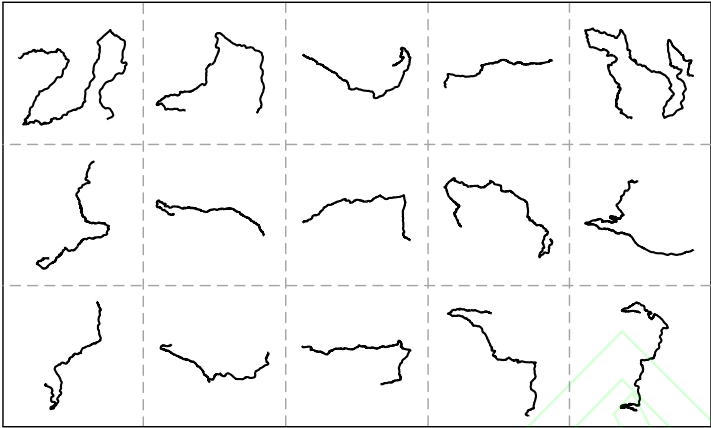


图 2 1:10K 河流数据（并非严格按比例显示）
Fig.2 Rivers at scale 1:10K (not shown to exact scale)

图 3 给出了其中 3 个河流要素在 7 种比例尺地图上的化简图形。篇幅所限，其余 12 条河流的化简图形从略。



图 3 3 条河流化简结果（并非严格按比例显示）
Fig.3 Simplification results of the three rivers (not shown to exact scale)

对于每一条河流要素，用式(8)分别计算原始河流(1: 10K)与化简后各比例尺地图上的对应河流图形之间的相似度。这样，可得到 15 条河流在 7 种比例尺变化下（1: 10K 图形可以与原始河流自身进行相似度计算）的 105 个相似度值。因为每个相似度（ Sim ）都与一个比例尺变化（ S ）相对应，所以可以用(S, Sim)组成 105 个坐标对。

考虑到相似度与比例尺变化之间的单调函数依赖关系，选取如下 6 个候选函数，用得到的 105 个坐标对作为拟合点，构建它们之间的函数关系：

$$1. \quad Sim = aS + b; \quad 2. \quad Sim = aS^2 + bS + c;$$

$$3. Sim = aS^3 + bS^2 + cS + d; \quad 4. Sim = a\ln(S + b) + c;$$

$$5. Sim = ae^{bS};$$

$$6. Sim = aS^b + b。$$

拟合结果如图 4 所示。其中，对数函数的 $R^2 = 0.92$ 最大，故选择对数函数作为比例尺变化与相似度的函数关系：

$$Sim = 2.122 - 0.275\ln(S + 60.764) \quad (10)$$

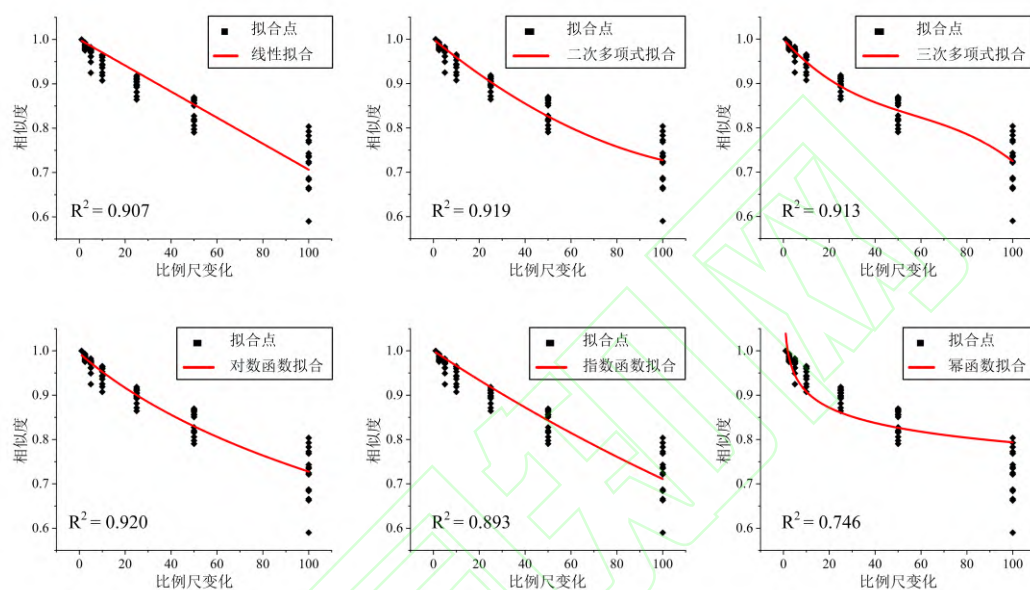


图 4 比例尺变化与相似度之间的函数关系拟合
Fig.4 Curve fitting of map scale change and similarity degree

4.2 Sim 与 ε 的函数关系推导方法

下面以某地 1:10K 地图水系要素的化简为例，推导相似度与 ε 的函数关系。如图 5，地图上有 16 条河流，相连的河流之间以红色圆点为分界，绿色方框为计算相似度过程中各河流要素对应的 MABR 矩形。

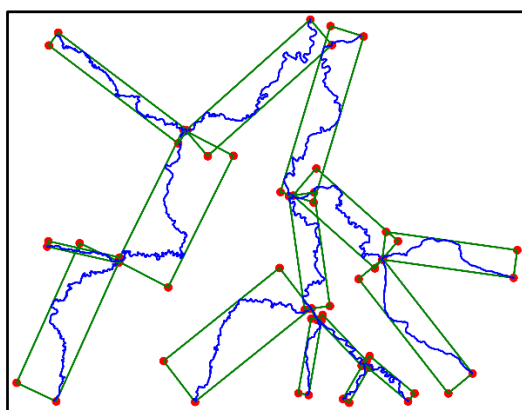


图 5 推导 ε 所用的河流数据
Fig.5 River data used in deducing ε

根据 DP 算法原理可以计算得到把所有河流均化简为直线时的 ε 值，此即为 ε 最大值。在本实验中，该最大值为 678.3m。为了得到可信度较高的函数关系，取 ε 步长为 0.5m，即 $\varepsilon = \{0.5m, 1.0m, 1.5m, \dots, 678.5m\}$ ，对图(5)中的河流数据运用 DP 算法进行 1356 次化简，得到原始数据在各个 ε 下的化简结果。

用式（9）计算原始河流数据与每个 ε 值下化简得到的河流数据的相似度 Sim ，由此得到 1356 个 (ε, Sim) 坐标对。对这些坐标对进行曲线拟合，结果如图 6。其中，对数函数的 $R^2 = 0.989$ 为最大者，所以其对应的函数被认为是本实验中 ε 与 Sim 的最佳函数关系：

$$Sim = -0.096\ln(\varepsilon) + 1.2097 \quad (11)$$

其反函数为：

$$\varepsilon = e^{12.601-10.471Sim} \quad (12)$$

一般来说，选取的实验数据不同， ε 与 Sim 的函数关系就不同。

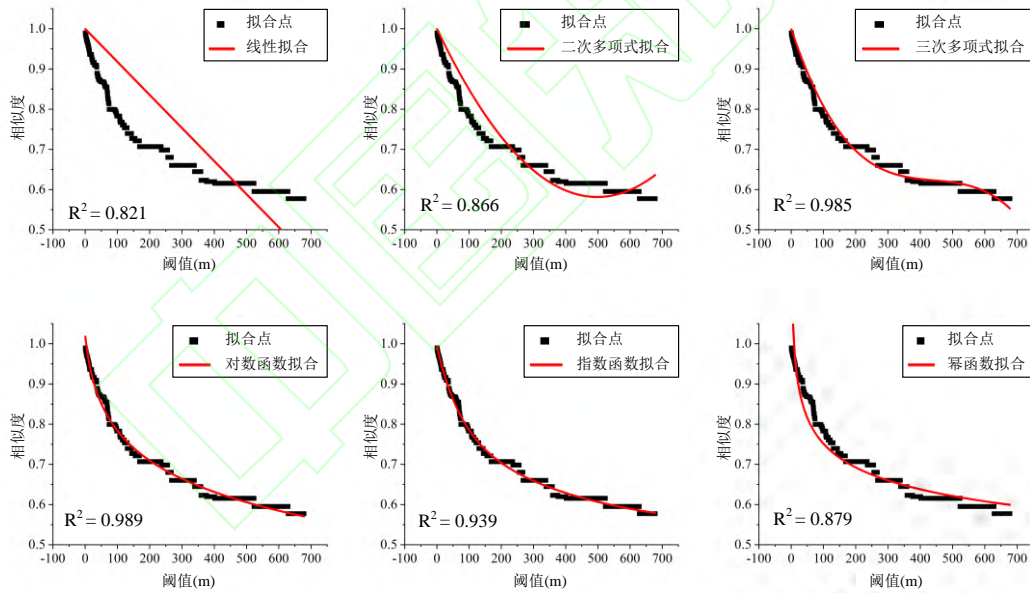


图 6 相似度与阈值的函数关系拟合
Fig.6 Curve fitting of similarity degree and ε

4.3 ε 与 S 的函数关系确定

结合式(10)、式(12)，可得到 S 与 ε 之间的函数关系：

$$\varepsilon = e^{12.601-10.471 \times (2.122-0.275\ln(S+60.764))} \quad (13)$$

对其化简得到：

$$\varepsilon = e^{0.275\ln(S+60.764)-9.618} \quad (14)$$

此即为图 5 数据的 ε 与 S 之间的函数关系。

5 实验与讨论

欲将图 5 中 1: 10K 河流数据用 DP 算法自动化简为 1: 25K、1: 50K、1: 100K、1: 250K、1: 500K、1: 1M 共 6 个比例尺数据，需用式(13)计算 ε ，结果见表 1。

表 1 距离阈值的计算
Tab.1 Calculation of ε

S	2.5	5	10	25	50	100
ε	10.215	11.425	14.104	24.533	51.247	149.802

图 7 是本文提出的自动 DP 算法化简结果（蓝色显示）、有经验制图员手工化简结果（红色显示）及二者的叠置对比图。用式(9)分别计算原始河流图形被自动 DP 算法和制图员化简到 1: 25K、1: 50K、1: 100K、1: 250K、1: 500K、1: 1M 比例尺河流时的相似度，结果如表 2 所示。

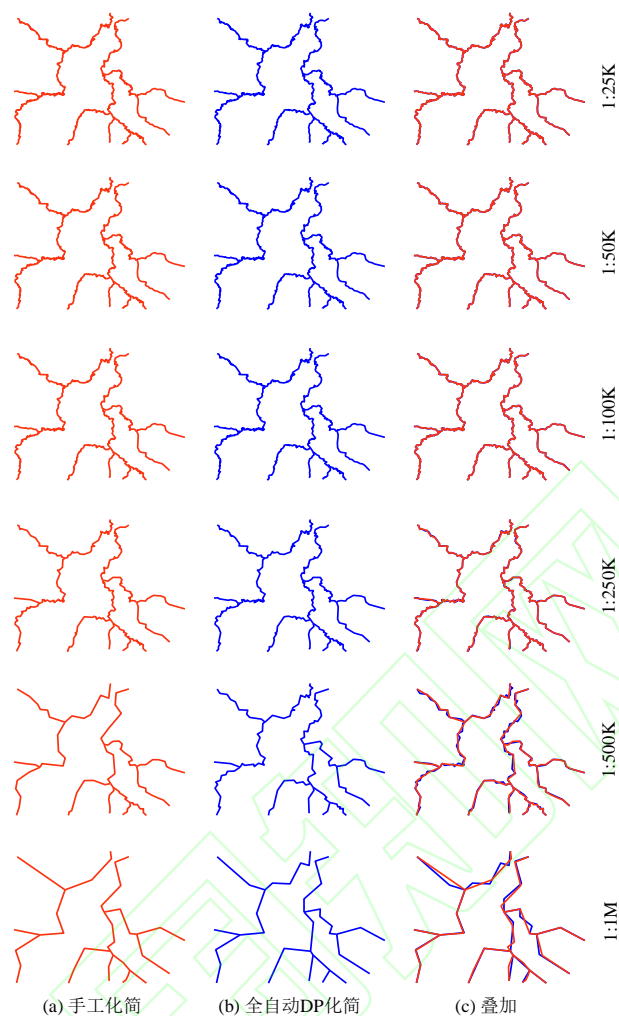


图 7 自动 DP 算法与手工化简结果的对比

Fig.7 Comparison of simplification results by the DP algorithm and that by cartographers

表 2 自动 DP 算法与人工方法对曲线化简结果的相似度比较

Tab.2 Comparison of similarity degrees of simplification results by the DP Algorithm and that by cartographers

比例尺	25K	50K	100K	250K	500K	1M
<i>Sim</i>	0.982	0.976	0.969	0.931	0.863	0.843

分析实验过程、曲线化简结果及 S 与 ε 之间的函数关系可得：

(1) 从图 7(a)-(c)可以看出，本文借助实验和公式推导给出的自动 DP 算法可以获得指定地理区域不同比例尺的水系要素的化简结果。

(2) 通过图 7(a)-(c)可视化对比发现，本文提出的自动 DP 算法得到的各个比例尺下的化简结果与制图员通过手工方式化简结果具有较高的相似度。

(3) 从表 2 中相似度定量比较发现，利用两种方法化简得到结果的相似度随着比例尺

变化增大而减小,同时,两种方法得到结果的相似度平均值为 0.927,相似度总体表现良好,说明本文算法具有较高的可靠性和智能化程度。

(4) 公式(14)是由 15 条河流在 7 种比例尺下的化简结果拟合而得,即一共用了 $7 \times 15 = 105$ 条曲线。根据作者的实验, $7 \times 14 = 98$ 条曲线就可使拟合得到的 S 与 ε 之间的函数关系基本趋于稳定,即曲线数量的继续增加虽然有助于提高函数的精度,但在曲线数量超过 98 条后对函数精度的提高幅度有限。

(5) 公式(14)适用于特定区域下的河流要素化简时 DP 算法中 ε 的确定。如果曲线地物的类型或地图综合区域的地理特点发生了变化,可以用类似的方法求得相应条件下的 S 与 ε 之间的函数关系。

6 结论

为了解决地图综合中常用的 DP 算法在执行前需要人工干预输入距离阈值 ε 而导致的非自动化问题,本文引入了多尺度曲线相似度的概念和计算方法,推导出了多尺度曲线相似度与比例尺变化的函数关系及相似度与 ε 的函数关系,进而以某地理区域的河流数据为例,推导出了 ε 的计算公式,实现了面向线状水系要素化简的 DP 算法自动化。各个比例尺下自动 DP 算法与制图员手工化简得到的图形相似度普遍较高,表明算法具有较高的可靠性。

虽然文中仅以河流数据为例进行了实验,但是该方法为其他线状地物(如道路、等高线、地类界)的全自动化简提供了一种有效思路。值得注意的是,本方法中 ε 计算公式的精度依赖于样本数据,样本越多、越典型,则公式的精度和可靠性越高。此外,针对不同的区域和不同的地物类型,需要推导不同的 ε 计算公式。因此,选取不同区域、不同模式、更大样本的河系进行实验,寻找更为普适的 ε 计算公式值得未来继续探索。

参考文献

- [1] R Weibel. A Typology of Constraints to Line Simplification[C]//Advance in GIS Research II ,Proceedings of 7th international symposium on spatial Data Handling. Lon-don: Taylor&Francis, 1997.
- [2] Jenkins J, Fleenor A, Dietz F. Moving beyond the Frame: Geovisualization of Landscape Change along the Southwestern Edge of Yosemite National Park[J]. *Journal of Geovisualization and Spatial Analysis*, 2019, 3(2): 9
- [3] Li Z L, Gong X Y, Jun C, et al. Functional Requirements of Systems for Visualization of Sustainable Development Goal (SDG) Indicators[J]. *Journal of Geovisualization and Spatial Analysis*, 2020, 4(1): 5
- [4] Mao B, Li B C. Graph-Based 3D Building Semantic Segmentation for Sustainability Analysis[J]. *Journal of Geovisualization and Spatial Analysis*, 2020, 4(1): 4
- [5] Larkey L B, Markman A B. Processes of Similarity Judgment[J]. *Cognitive Science*, 2005, 29(6): 1061-1076
- [6] Visser H, de Nijs T. The Map Comparison Kit[J]. *Environmental Modelling & Software*, 2006, 21(3): 346-358
- [7] Li B N, Fonseca F. TDD: A Comprehensive Model for Qualitative Spatial Similarity Assessment[J]. *Spatial Cognition & Computation*, 2006, 6(1): 31-62
- [8] Rodriguez M A, Egenhofer M J. Determining Semantic Similarity among Entity Classes from Different Ontologies[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2003, 15(2): 442-456
- [9] Liu Pengcheng, Xiao Tianyuan, Xiao Jia, et al. A Head-Tail Information Break Method Oriented to Multi-Scale Representation of Polyline[J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(7): 921-933 (刘鹏程, 肖天元, 肖佳, 等. 曲线多尺度表达的 Head-Tail 信息量分割法 [J]. 测绘学报, 2020, 49(7): 921-933)
- [10] Du Jiawei, Wu Fang, Li Jinghan, et al. A Progressive Simplification Method for the Estuary Coastline[J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(4): 547-556 (杜佳威, 武芳, 李靖涵, 等. 一种河口湾海岸线渐进化简方法[J]. 测绘学报, 2018, 47(4): 547-556)

- [11] Ramer U. An Iterative Procedure for the Polygonal Approximation of Plane Curves[J]. *Computer Graphics and Image Processing*, 1972, 1(3): 244-256
- [12] Douglas D H, Peucker T K. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Caricature[J]. *Cartographica: the International Journal for Geographic Information and Geovisualization*, 1973, 10(2): 112-122
- [13] Hershberger J, Snoeyink J. Speeding up the Douglas–Peucker line-simplification algorithm[C]// Proceedings of the 5th Symposium on Data Handling, 1992: 134–143
- [14] Yan Haowen, Wang Jiayao. A Generic Algorithm for Point Cluster Generalization Based on Voronoi Diagrams[J]. *Journal of Image and Graphics*, 2005, 10(5): 633-636 (闫浩文, 王家耀. 基于 Voronoi 图的点群目标普适综合算法[J]. 中国图象图形学报, 2005, 10(5): 633-636)
- [15] Li Shibao, Chen Tong, Liu Jianhang, et al. A Road Curve Simplification Algorithm Based on Intersection Point[J]. *Engineering of Surveying and Mapping*, 2017, 26(7): 1-4 (李世宝, 陈通, 刘建航, 等. 基于交叉点的道路曲线化简算法研究[J]. 测绘工程, 2017, 26(7): 1-4)
- [16] Yu Jing, Chen Gang, Zhang Xiao, et al. An Improved Douglas-Peucker Algorithm Oriented to Natural Shoreline Simplification[J]. *Science of Surveying and Mapping*, 2015, 40(4): 23-27 (于靖, 陈刚, 张笑, 等. 面向自然岸线抽稀的改进道格拉斯—普克算法[J]. 测绘科学, 2015, 40(4): 23-27)
- [17] Yan H W. Quantitative Relations between Spatial Similarity Degree and Map Scale Change of Individual Linear Objects in Multi-Scale Map Spaces[J]. *Geocarto International*, 2015, 30(4): 472-482
- [18] Yan H W, Shen Y Z, Li J. Approach to Calculating Spatial Similarity Degrees of the Same River Basin Networks on Multi-Scale Maps[J]. *Geocarto International*, 2016, 31(7): 765-782
- [19] Chen Z L, Ye W. The Precise Representation Model of Topological Relations of Complex Planar Objects[J]. *Journal of Geodesy and Geoinformation Science*, 2019, 2(3): 18-30
- [20] Huang B H, Zhong W, Zhai R J, et al. Hierarchical Area Partitioning Method of Urban Road Networks Matching[J]. *Journal of Geodesy and Geoinformation Science*, 2019, 2(3): 55-67
- [21] Toussaint G T. A comparison of rhythmic dissimilarity measures[J]. *FORMA*. 2006, 21(2): 129–149

- [22] Rosso R, Bacchi B, La Barbera P. Fractal Relation of Mainstream Length to Catchment Area in River Networks[J]. *Water Resources Research*, 1991, 27(3): 381-387
- [23] Power C, Simms A, White R. Hierarchical Fuzzy Pattern Matching for the Regional Comparison of Land Use Maps[J]. *International Journal of Geographical Information Science*, 2001, 15(1): 77-100
- [24] Chen Qingyan, Liang Dan, Xu Wenbing, et al. Study of the Linear Object Similarity Measure Based on Hausdorff Distance[J]. *Science of Surveying and Mapping*, 2016, 41(8): 14-18 (陈青燕, 梁丹, 徐文兵, 等. 一种线目标豪斯多夫相似距离度量指标[J]. 测绘科学, 2016, 41(8): 14-18)
- [25] Deng Min, Niu Shulian, Li Zhilin. A Generalized Hausdorff Distance for Spatial Objects in GIS[J]. *Geomatics and Information Science of Wuhan University*, 2007, 32(7): 641-645 (邓敏, 钮沐联, 李志林. GIS 空间目标的广义 Hausdorff 距离模型[J]. 武汉大学学报·信息科学版, 2007, 32(7): 641-645)
- [26] Cheng Mianmian, Sun Qun, Li Shaomei, et al. Generalized Hausdorff Distance of Multi-Scale Point Group and Its Application in Similarity Measurement[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(6): 885-891 (程绵绵, 孙群, 李少梅, 等. 多尺度点群广义 Hausdorff 距离及在相似性度量中的应用[J]. 武汉大学学报·信息科学版, 2019, 44(6): 885-891)
- [27] Huttenlocher D P, Klanderman G A, Rucklidge W J. Comparing Images Using the Hausdorff Distance[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993, 15(9): 850-863

Approach to automating the DP Algorithm—taking river simplification as an example

YAN Haowen^{1,2,3}, ZHANG Xingang¹, LU Xiaomin^{1,2,3}, LI Pengbo¹

1. Department of GIS, Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China

2. National-local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou 730070, China

Abstract: Objectives: Curve simplification is of importance in automated map generalization; nevertheless, the Douglas-Peucker Algorithm (i.e., the DP Algorithm) popularly used in map generalization is not automatic, because a key parameter called distance tolerance (i.e., ε) must be given by experienced cartographers and need to be input before execution of the algorithm. **Methods:** To solve the problem, this paper proposed a method to automatically calculate ε and by which the automation of the DP Algorithm is achieved. The method consists of the following steps: (1) A formula is constructed by the Hausdorff Distance for calculating the similarity degree (Sim) between a curve at a larger scale and its simplified counterpart at a smaller scale. (2) 15 linear rivers are selected, and each of them are manually simplified to get their counterparts at seven different scales. The Sim of each original river and each of its simplified counterpart at a smaller scale can be obtained using the formula constructed by the Hausdorff Distance, and $15 \times 7 = 105$ coordinate pairs consisting of (S, Sim) can be got, and a function between Sim and S are constructed by the curve fitting using the coordinates. (3) In the meanwhile, the 15 rivers are simplified using a number of ε , and the Sim of each original river and each of its simplified counterpart at a smaller scale can be calculated using the formula constructed by the Hausdorff Distance. In this way, a number of coordinate pairs (ε, Sim) are got, and a function between ε and Sim is constructed by the curve fitting. (4) By the function between Sim and S and that between ε and Sim , a formula between ε and S can be deducted. Using the formula ε can be calculated automatically, because in a map generalization task S is usually known. After this step, automation of the DP Algorithm is achieved. **Results:** The experiments have shown that (1) the proposed DP Algorithm can automatically simplify the rivers in a specific geographical area to get the results at different scales; and (2) the resulting river curves generated by the proposed DP Algorithm have a high degree of similarity with the ones made by experienced cartographers. Their average similarity degree is 0.927. **Conclusion:** The proposed DP algorithm can simplify curve features on maps automatically, and the results are highly intelligent and credible. Although only river data is tested in this paper, the principle of the proposed method can be extended to other linear features on maps. Our future work will be on improving the accuracy of the proposed DP Algorithm using more river data so that the algorithm can be used in practical map generalization engineering.

Keywords: Curve simplification; map generalization; Douglas-Peucker Algorithm; similarity degrees; river data.

First author: YAN Haowen, PhD, professor, majors in map generalization and spatial analysis. E-mail: yanhw@mail.lzjtu.cn.

Foundation support: The National Natural Science Foundation of China (41930101).

网络首发:

标题: DP算法自动实现方法—以河流化简为例

作者: 闫浩文, 张鑫港, 禄小敏, 李蓬勃

DOI: 10.13203/j.whugis20210412

收稿日期: 2022-07-13

引用格式:

闫浩文, 张鑫港, 禄小敏, 等. DP算法自动实现方法—以河流化简为例[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210412 (YAN Haowen, ZHANG Xingang, LU Xiaomin, et al. Approach to Automating the DP Algorithm—Taking River Simplification as an Example[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210412)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

多尺度点群广义 Hausdorff 距离及在相似性度量中的应用

程绵绵, 孙群, 李少梅, 徐立

武汉大学学报·信息科学版, 2019, 44(6): 885-891

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20170305>