

基于空间数据挖掘的分区异步元胞 自动机模型研究

柯新利^{1),2),3)} 边馥苓¹⁾

¹⁾(武汉大学空间信息与数字工程研究中心, 武汉 430079) ²⁾(中国科学院地理科学与资源研究所, 北京 100101)

³⁾(咸宁学院资源与环境科学学院, 咸宁 437000)

摘要 传统的元胞自动机模型采用统一的转换规则和相同的演化速率进行演化,忽略了地理现象演变的时空差异性:演化规律的空间异质性和演化速率的空间差异性。针对这一问题,提出了基于空间数据挖掘的分区异步元胞自动机模型,采用双约束空间聚类的方法对元胞空间进行分区,用分区转换规则替代统一转换规则可以体现地理现象演化规律的空间差异性;采用标准格网划分的方法求取异步元胞演化速率,用异步演化速率替代同步演化速率可以体现地理现象演化速率的空间差异性。以杭州市土地利用变化为例对基于空间数据挖掘的分区异步元胞自动机模型进行了实证研究,结果表明:与传统的元胞自动机模型相比,基于空间数据挖掘的分区异步元胞自动机模型具有较高的模拟精度,并且适用于较大区域较长时间段地理现象的动态变化模拟。基于空间数据挖掘的分区异步元胞自动机模型是地理元胞自动机研究的新视角,它将地理现象演变的时空异质性和时间差异性引入到地理元胞自动机模型中,使模型对地理过程的模拟更接近实际地理过程。然而,由于有关分区异步的元胞自动机模型还处于尝试性研究阶段,在元胞空间分区方法、双约束空间聚类算法中权重的确定方法、元胞演化速率的获取方法、元胞转换规则的获取方法、模拟精度评估以及分区异步元胞自动机模型在较大区域较长时间的地理现象模拟中的应用等方面有待进一步的研究与探讨。

关键词 元胞自动机 空间数据挖掘 分区 异步

中图法分类号: TP23 **文献标志码:** A **文章编号:** 1006-8961(2010)06-921-10

A Partitioned & Asynchronous CA Based on Spatial Data Mining

KE Xinli^{1),2),3)}, BIAN Fuling¹⁾

¹⁾(Spatial Information & Digital Engineer Research Center of Wuhan University, Wuhan 430079)

²⁾(Institute of Geographic Science & Natural Resources Research, Chinese Academy of Sciences, Beijing 100101)

³⁾(School of Resources & Environment Science, Xianning College, Xianning 437000)

Abstract In traditional CA, general transfer rule and same speed are used to drive models. In this kind of CA, spatial-temp differences of geographical phenomenon, both transfer rules and transfer speed, are ignored. To solve this problem, a partitioned & asynchronous CA based on spatial data mining is given in this paper. In this model, cell space is departed by dual-constraint spatial cluster and general transfer rule is replaced by partition transfer rule, asynchronous transfer speed is calculated using general grid and synchronous transfer speed is replaced by asynchronous transfer speed. Spatial differences of geographical transfer rule and geographical transfer speed are taken into account in this kind of models. Taking land use change in Hangzhou as a case, partitioned & asynchronous CA based on spatial data mining are implemented in land use. Result shows: comparing with traditional CA, partitioned & asynchronous CA based on spatial data mining can get more accurate simulation result, and this kind of models can be used to simulate geographical phenomenon in a larger area for a comparatively longer time. A new viewpoint of GeoCA is given in partitioned & asynchronous CA. In this kind of Model,

基金项目:国家自然科学基金资助项目(70873118);国家科技支撑计划(2006BAB15B02);湖北省自然科学基金(2009CDB342);湖北省教育厅人文社会科学项目(2010Q130)

收稿日期:2009-04-08;**改回日期:**2009-09-01

第一作者简介:柯新利(1977—),男,讲师。2009年6月于武汉大学获得工学博士学位,现为中国科学院地理科学与资源研究所博士后。主要研究方向为GIS空间分析、土地系统动态建模。E-mail: kevin.kxl@gmail.com

spatial differences and temporal diversities are taken account into GeoCA, it makes simulations with this kind of model much closer to actuality. Yet theories and methods of partitioned & asynchronous CA is still in tentative research stage, there are many problems, such as divided methods of cell space, methods of calculating weight of dual-constraint spatial cluster, calculation of cell transfer speed, getting of transfer rules, evaluation of CA result accuracy, and applications of partitioned & asynchronous CA in simulating land use at larger area and during longer time and etc. should be discussed and analyzed.

Keywords CA, spatial data mining, partition, asynchronous

0 引言

元胞自动机(CA)具有模拟复杂系统时空演化过程的能力,它“自下而上”的研究思路充分体现了复杂系统局部的个体行为产生全局、有秩序模式的理念,体现了“复杂系统来自简单子系统的相互作用”这一复杂性科学的精髓,使得它在地理学的研究中具有天然优势^[1]。近年来,越来越多的学者利用元胞自动机来模拟复杂的地理现象,并且取得了许多有意义的研究成果^[2-9]。这些研究表明,元胞自动机与GIS结合具有复杂地理现象和过程的模拟和预测能力,是现有GIS时空分析功能的发展和丰富,可以克服GIS在动态空间建模与分析方面的局限性。

当前的有关地理元胞自动机模型的研究主要集中在元胞转换规则的获取方面。Jenerette在CA模型的转换规则上,引入了遗传算法^[10]。Liu和Andersson将马尔可夫随机场与元胞自动机模型相结合来模拟城市增长情况^[11]。Takeshi和Tetsuya将回归分析和差别分析与元胞自动机模型相结合,对东京郊区的城市扩展情况进行了模拟^[12]。Liu和Stuart将模糊集函数和CA模型相结合,对悉尼市从1971年到1996年的城市发展情况进行了很好的模拟^[13]。Claudia等人用随机元胞自动机模型和Bayes经验方法相结合,用来模拟预测巴西西部城市Bauru的扩展情况^[14]。国内也有很多学者对元胞自动机转换规则的获取进行了研究。黎夏和叶嘉安将人工神经网络引入到元胞自动机模型中,构建了ANN-CA模型^[15]。刘小平等人提出了基于蚁群智能自动获取转换规则的方法,基于此构建了基于蚁群智能算法的地理元胞自动机(ACO-CA)^[16]。胡茂桂等人应用BP神经网络与元胞自动机相结合自动挖掘元胞自动机模型的转换规则^[17]。杨青生等人运用多智能体(Agent)和元胞自动机结合来模拟城市用地扩张^[18]。

然而,目前大多数元胞自动机模型在模拟地理

现象的动态变化时,存在着一些局限性。一方面,大多数元胞自动机模型采用统一的转换规则驱动元胞自动机模型中所有元胞的演化来模拟地理现象的动态演化,忽视了被模拟地理现象及其相关因素的空间异质性,因而从一定程度上限制了地理元胞自动机模型模拟精度的提高;另一方面,大多数元胞自动机模型在整个元胞空间内所有的元胞都以同样的速率进行演化,忽视了被模拟地理现象演化速率的空间差异性,与实际的地理现象动态变化规律不符,因而成为地理元胞自动机模型模拟精度提高的另一个瓶颈。针对当前元胞自动机模型存在的上述两方面的问题,提出了基于空间数据挖掘的分区异步元胞自动机模型,利用空间数据挖掘的方法对元胞空间进行分区并对各个分区分别求取元胞转换规则,结合不同的演化速率驱动元胞自动机模型进行地理现象的模拟。

1 模型构建

1.1 模型原理

1.1.1 总体框架

以土地利用变化模拟为例,基于空间数据挖掘的分区异步元胞自动机模型的原理如图1所示。将研究期初的土地利用数据和研究期末的土地利用数据进行叠加,得到土地利用变化数据。对土地利用变化数据和其他影响土地利用变化的数据一起进行空间聚类分析,聚类结果中的每一个类(簇)就对应元胞空间的一个分区。另一方面,对期末的土地利用数据和影响土地利用变化的各影响因子进行采样,并将采样得到的土地利用样本数据和聚类结果一同输入到决策树单元进行处理得到各个分区的转换规则。与此同时,异步演化速率计算单元将元胞空间划分成很多速率网格,并且根据土地利用变化数据计算得到每一个速率网格的元胞演化速率。在此基础上,将得到的元胞演化速率和各个分区的元胞转换规则一起输入到元胞自动机模型进行研究区域的土地利用变化模拟,将模拟的结果和期末土地

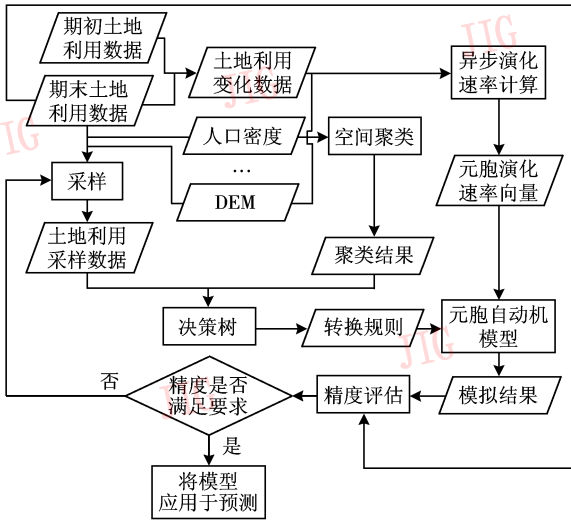


图 1 基于空间数据挖掘的分区异步元胞自动机模型原理图

Fig. 1 Working flow of partitioned asynchronous CA based on spatial data mining

利用数据进行对比,可以得出元胞自动机模型的模拟精度。在此基础上,对元胞自动机模型的模拟精度进行评估。若精度能够达到要求,则可用该元胞自动机模型对未来的土地利用变化情况进行模拟和预测;否则,需要对期末的土地利用数据及各影响因子进行重采样,重新获取元胞转换规则驱动元胞自动机进行土地利用变化模拟,直到精度达到要求。

基于空间数据挖掘的分区异步元胞自动机模型具有两个方面的特点:一是根据地理现象的空间异质性,采用双约束空间聚类的方法将整个元胞空间划分成几个相对匀质的区域,然后对每个分区分别获取转换规则;二是根据地理过程演化的速率不同,采用标准网格将整个元胞空间划分为若干个子区域,分别对每一个子区域求取元胞演化的速率,利用各子区域的演化速率来驱动元胞自动机模型按特定的演化速率进行演化。根据地理现象的空间异质性原理而提出的分区保证了在各个子区域内的元胞具有相对一致的元胞转换规则,根据地理演变速率的空间差异性提出的异步则保证了在同一个速率格网内的元胞具有相对一致的演化速率。

1.1.1.2 元胞空间分区

元胞空间就是元胞所分布的空间网点集合。针对某一项具体的应用而言,元胞空间就是指整个研究区域的范围。由于地理现象广泛存在着空间异质性,所以在元胞自动机模型中,转换规则的空间异质

性也因地理现象的空间异质性而客观存在。为了减小不一致的元胞转换规则带来的元胞自动机模型模拟误差,有必要对元胞空间进行分区,使得每一个区域具有相对一致的元胞转换规则。

在对元胞空间进行分区时,不仅要保证同一分区的元胞在空间位置上相互邻近,而且要保证同一分区的元胞在非空间属性方面彼此接近。基于传统的空间聚类算法对元胞空间进行分区聚类的依据仅是元胞相互之间的空间位置,没有考虑元胞的属性数据。因此,引入双约束空间聚类算法对元胞空间进行分区。

定义 双约束空间聚类:给定一空间对象集 $S = \{O_1, O_2, \dots, O_n\}$,任意对象 $O_i \in S$ 有空间域和非空间域两个属性域,双约束空间聚类就是将对象集 S 划分成不同的簇,满足每个簇在空间域上形成紧凑的区域而在非空间域上的相似性最大化^[19]。

双约束空间聚类和传统空间聚类的本质区别在于对聚类统计量(即距离)的定义不同。传统的空间聚类更加关注空间对象的空间距离,而双约束空间聚类不仅关注空间对象的空间距离,同时也关注空间对象之间的属性距离。因此,可以通过重新定义距离来实现双约束空间聚类。为了克服空间几何距离不足以进行双约束空间聚类的局限,采用广义 Euclid 距离替代空间几何距离作为双约束空间聚类的聚类统计量,广义 Euclid 距离的定义如下^[20]:

$$D_{ij} = w_p \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} + w_a \sqrt{\sum_{k=1}^m w_k (z_{ik} - z_{jk})^2} \quad (1)$$

$$w_p + w_a = 1$$

$$\sum_{k=1}^m w_k = 1$$

式中, D_{ij} 为点 i 和点 j 之间的广义 Euclid 距离; (x_i, y_i) 和 (x_j, y_j) 分别表示点 i 和点 j 的空间坐标; z_{ik}, z_{jk} 分别表示点 i 和点 j 的第 k 个属性值; m 为点群的属性特征数目; w_p, w_a 为在广义 Euclid 距离中空间距离和属性距离的权重; w_k 为属性距离计算中各个属性的权重。当 $w_p = 1, w_a = 0$ 时, D_{ij} 即为空间距离,空间聚类即变为单纯的空间几何距离聚类;当 $w_p = 0, w_a = 1$ 时, D_{ij} 即为属性距离,空间聚类变为单纯的属性聚类。

1.1.1.3 元胞演化速率获取

元胞演化速率是指异步元胞自动机模型中,元胞空间中每一个元胞在模拟过程中演化的快慢程

度。由于地理过程的演变速率存在空间差异性,作为地理过程模拟工具的地理元胞自动机模型的元胞演化速率也存在空间差异性,元胞速率正是这种差异性的量化表达。

不难发现,在同一个时间段内,对于大小相同的子区域而言,状态发生变化的栅格单元越多则表明地理现象的演化速率越快,反之亦然。

在求取元胞自动机模型的元胞转换速率时,先将整个研究区域划分为若干个子区域,然后得到不同子区域的元胞演化速率。为了保证各子区域大小的一致性,选择采用标准格网划分的方法将研究区域划分为很多标准格网,得到一个速率格网。

为了使元胞转换速率反映地理过程变化的快慢,采用下式来求取元胞转换速率:

$$v_i = \frac{c_i}{s_i} \quad (2)$$

式中, v_i 为速率格网 i 中的元胞演化速率, c_i 为速度格网 i 中发生变化的栅格(元胞)数目, s_i 为速率格网 i 中所有非空栅格(元胞)的数目。

式(2)中的 v_i 虽然能反映速率格网 i 中元胞演变的快慢,但是 v_i 很难与元胞自动机模型结合起来驱动元胞自动机模型的演变,有必要对 v_i 进行改进,使它能与元胞自动机模型结合起来,控制元胞自动机模型中某个或某些元胞的演化速率。这里,用元胞演化间隔的概念来替代元胞演化速率 v_i ,可以

用式(3)来计算元胞演化间隔:

$$Interval_i = \left\lceil \frac{1}{v_i} \right\rceil \quad (3)$$

式中, $Interval_i$ 即为速率格网 v_i 中的元胞演化间隔。

值得注意的是,由于 $0 \leq v_i \leq 1$, 所以 $1 \leq Interval_i \leq \infty$, 如果直接采用 $Interval_i$ 来控制元胞自动机模型的演化则有可能出现有些元胞永远也不可能发生演化的情况;此外,除去 $Interval_i = \infty$ 的情况后,也有一些元胞的演化间隔会远大于其他元胞的演化间隔。为了避免这种情况的发生,这里定义了最大元胞演化间隔 $MaxI$, 然后将所有元胞的演化间隔都控制在 $[1, MaxI]$ 区间。为了消除一些异常值的影响,将演化间隔大于 $MaxI$ 的情况全部用 $MaxI$ 代替。可以用式(4)表示:

$$Interval_i = \begin{cases} Interval_i & Interval_i \leq MaxI \\ MaxI & Interval_i > MaxI \end{cases} \quad (4)$$

在与元胞自动机模型结合时,先计算元胞空间每一个元胞的演化间隔,在演化时,可以判断当前元胞自动机模型运行的次数是否为该元胞演化间隔 $Interval_i$ 的整数倍,若是,则结合该元胞的转换规则判断该元胞是否发生演变;否则,该元胞不发生演化。

1.2 模型实现

基于空间数据挖掘的分区异步元胞自动机模型总体框架如图 2 所示。模型主要由 5 个模块组成: GIS 模块、双约束空间聚类模块、C5.0 决策树模块、异步演化速度模拟、元胞自动机模型。

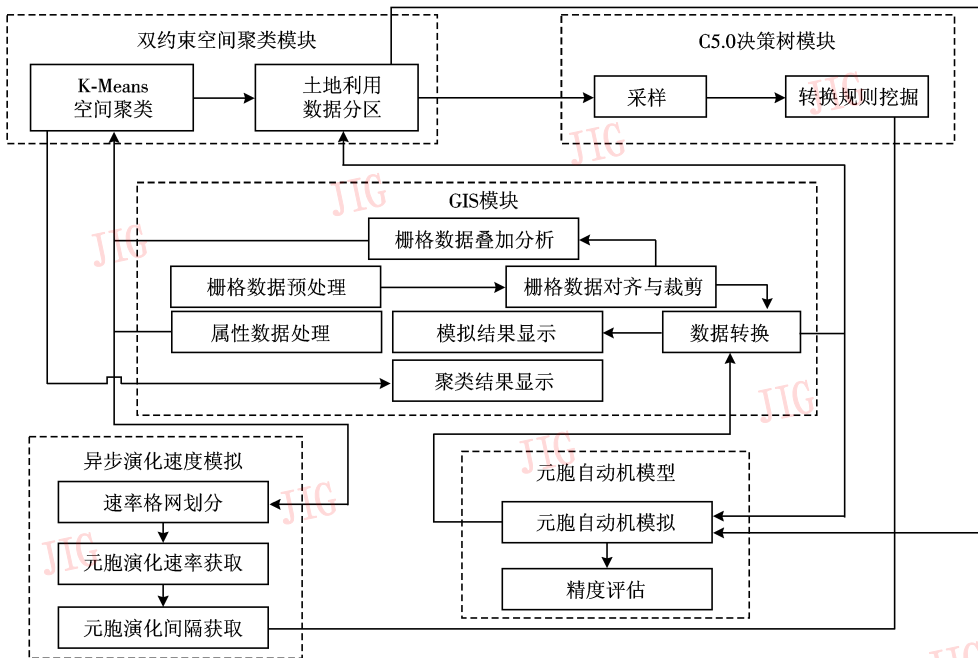


图 2 基于空间数据挖掘的分区异步元胞自动机模型总体构架

Fig. 2 Structure of partitioned & asynchronous CA based on spatial data mining

异步演化速率模块和元胞自动机模块。其中, GIS 模块的主要功能是 GIS 数据处理、分析、转换和显示等, 主要可以实现栅格数据的叠加分析、栅格数据对齐与裁剪、GIS 数据预处理、数据转换、结果显示等; 双约束空间聚类模块主要实现双约束空间聚类以及利用聚类结果实现研究区域(元胞空间)的分区; C5.0 决策树模块主要实现元胞转换规则的挖掘, 主要包括数据采样和 C5.0 决策树分析两个部分; 异步演化速率模块的主要功能是获取元胞演化速率和元胞演化间隔, 主要包括元胞自动机元胞演化速率格网划分、元胞演化速率的获取和元胞演化间隔的获取等; 元胞自动机模型部分的主要功能是根据 C5.0 决策树模块获取元胞转换规则和异步演化速率模块获取的元胞演化间隔进行元胞自动机模型模拟, 并对模拟结果进行精度评估。

2 模型应用

2.1 研究区域与数据处理

杭州市是以上海为中心的长江三角洲南翼中心

城市, 位于杭嘉湖平原和浙西丘陵山地的连接地带。全市面积 16 596 km², 其中市辖区 3 068 km²。辖上城、下城、拱墅、江干、西湖、滨江、萧山、余杭 8 个区, 临安、富阳、建德 3 个县级市, 桐庐、淳安 2 个县。本文选择的研究范围如图 3 所示, 主要包括杭州市城市规划的一主三副区域, 即主城、下沙城、江南城以及临平城, 体现了杭州市未来发展的组团式格局(图 3)。

本文研究的基础数据包括研究区域 2000 年及 2005 年两个时相的 Landsat TM/ETM + 遥感影像、交通数据、地形图、人口数据、水域数据。具体处理过程如下: 利用 ERDAS Imagine 8.7 软件对遥感影像进行几何校正、辐射校正、边界裁剪、监督分类及人工目视解译、精度检验, 生成土地利用图层; 利用 ArcGIS9.2 对地形图进行数字化并生成 DEM 数据, 利用 ArcGIS9.2 对交通数据、水域数据和人口数据进行处理, 生成交通图层、水域图层和人口密度图层, 在此基础上生成道路距离栅格、铁路距离栅格及水域距离栅格等距离栅格数据。

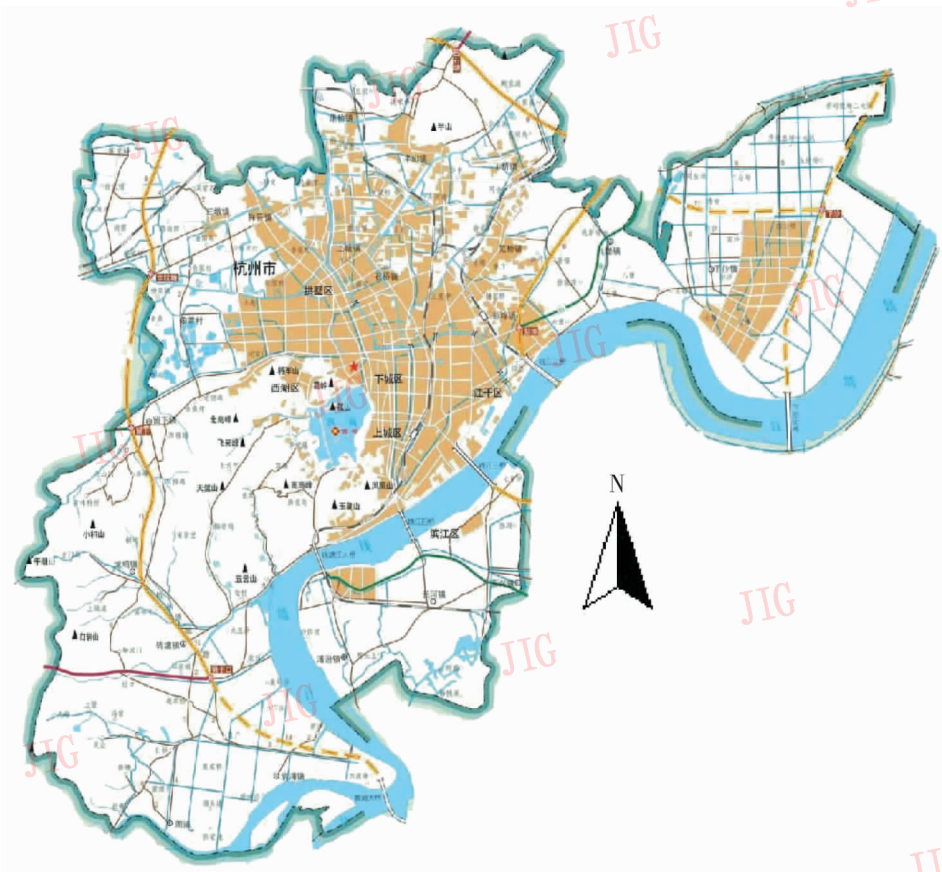


图3 研究区域范围

Fig.3 Research area

2.2 模拟结果

2.2.1 双约束聚类结果

根据前面提出的双约束空间聚类方法,首先对杭州市 2000 年和 2005 年的土地利用数据进行叠加运算,得到杭州市 2000 年到 2005 年的土地利用变化图层。然后将杭州市的土地利用变化图层与杭州市的人口密度图、杭州市交通数据集(包括铁路、公路和环线)、杭州市水域数据集等一起输入到双约束空间聚类模型进行空间聚类,得到如图 4 所示的聚类结果。

值得注意的是:1)本文中采用距离栅格来表示

交通数据集和水域数据集等,即用每个栅格离交通线路或水域的距离来表示,这些数据集本身已经包括了空间信息,但由于单纯的元胞与元胞之间的距离并不能很好地反映每一个元胞与交通线路或水域之间的关系,因此,本文仍然把这些包含了空间信息的数据集当成属性数据集来对待,以更好地反映每一个元胞与交通线路和水域的关系,得出更为合理的聚类效果;2)在采用双约束空间聚类对元胞空间进行分区时,权重 w_p 和 w_a 的确定十分重要,直接关系到元胞分区的效果,本文根据经验采用逐步试算的方法确定权重 w_p 和 w_a 以获得较好的分区效果。

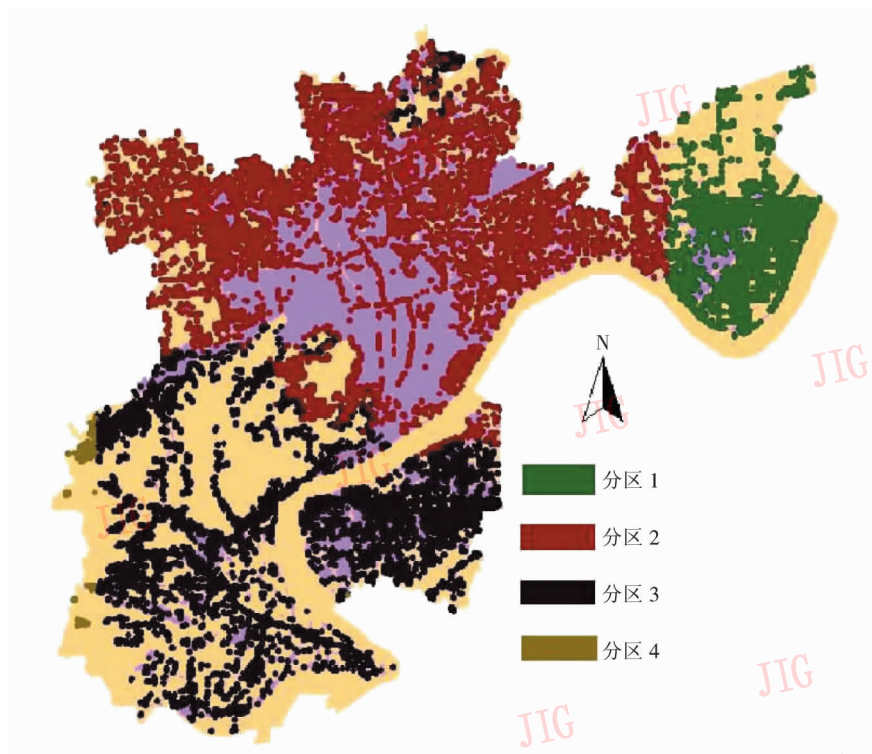


图 4 双约束空间聚类结果

Fig. 4 Result of dual-constraint spatial cluster

与传统的空间聚类方法相比,双约束空间聚类可以较好地反映元胞相互之间以及元胞与相关要素层之间的关系,如图 5 所示。

由图 5 可以看出,双约束空间聚类可以较好地反应元胞相互之间以及元胞与相关要素层之间的关系,双约束空间聚类所得到的分区(簇)内的各空间对象之间不仅在空间上相互邻近,而且簇内具有相对一致的属性信息。因此,双约束空间聚类可以较好地反映由空间属性和非空间属性共同作用产生的空间分异规律,可以获得更为科学和合理的分区(簇)。

2.2.2 异步演化速率计算结果

采用提出的元胞演化间隔的求取方法,对杭州市 2000 年到 2005 年的土地利用数据进行计算,得到了杭州市 2000 年到 2005 年的土地利用演变间隔的栅格图,如图 6 所示。图中显示的土地利用演变间隔栅格图是经过规整化后的栅格图。图中速率格网的大小为 10×10 个栅格单元(元胞),最大演化间隔为 100。

2.2.3 CA 模拟结果

以杭州市 2000 年的土地利用情况为起点,利用基于空间数据挖掘的分区异步元胞自动机模型

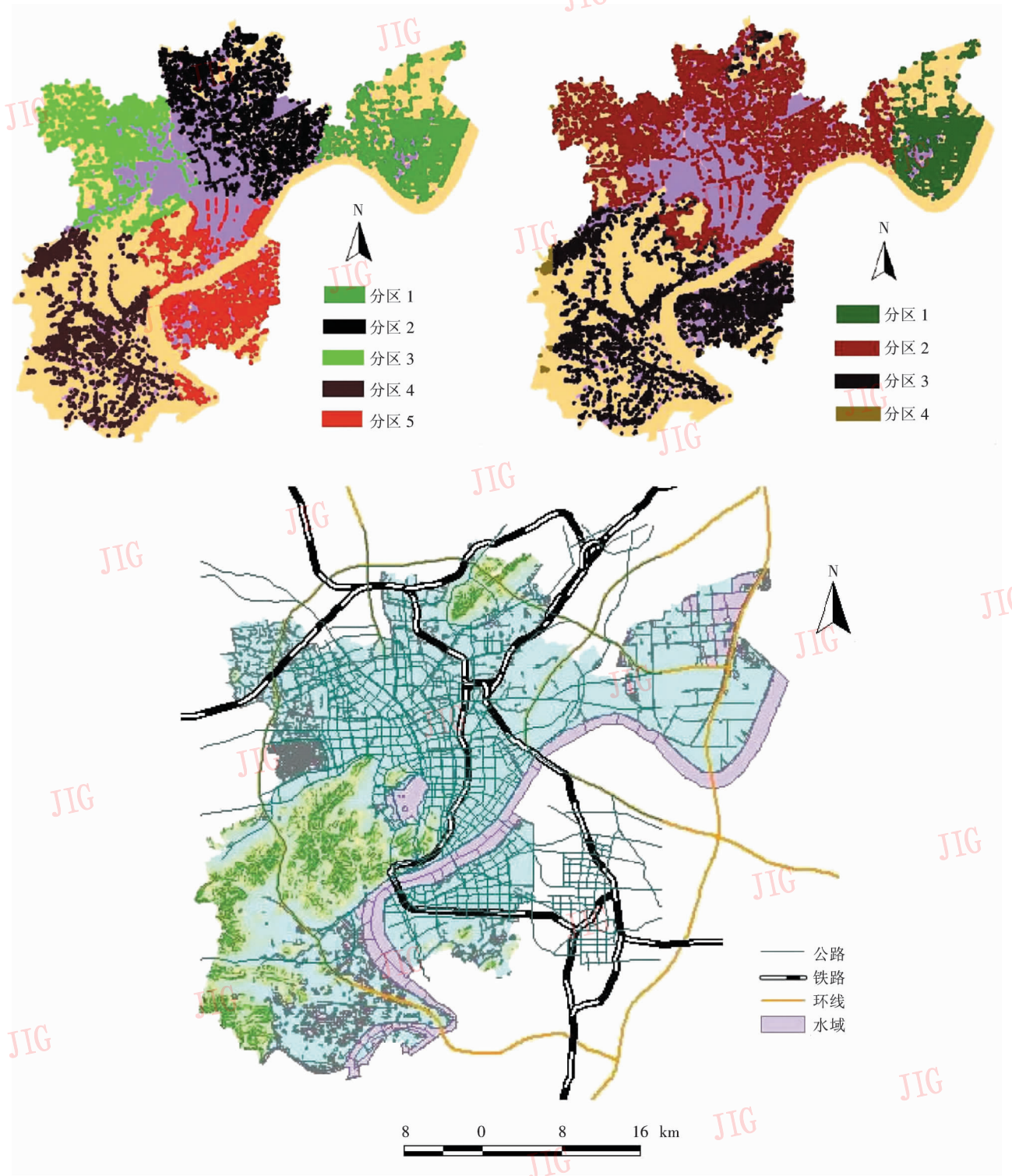


图 5 传统空间聚类与双约束空间聚类的结果对比

Fig. 5 Comparing result of traditional spatial cluster and dual-constraint spatial cluster

对杭州市 2005 年的土地利用情况进行了模拟,得出如图 7 所示的结果。

从模拟结果可以看出,与传统的元胞自动机模型相比,基于双约束空间聚类的异步元胞自动机模

型的模拟结果与杭州市 2005 年的土地利用情况更加接近。这就说明了采用分区演化规则和异步演化速率驱动元胞自动机模型进行土地利用变化模拟更加符合实际情况,可以获得更高的模拟精度。

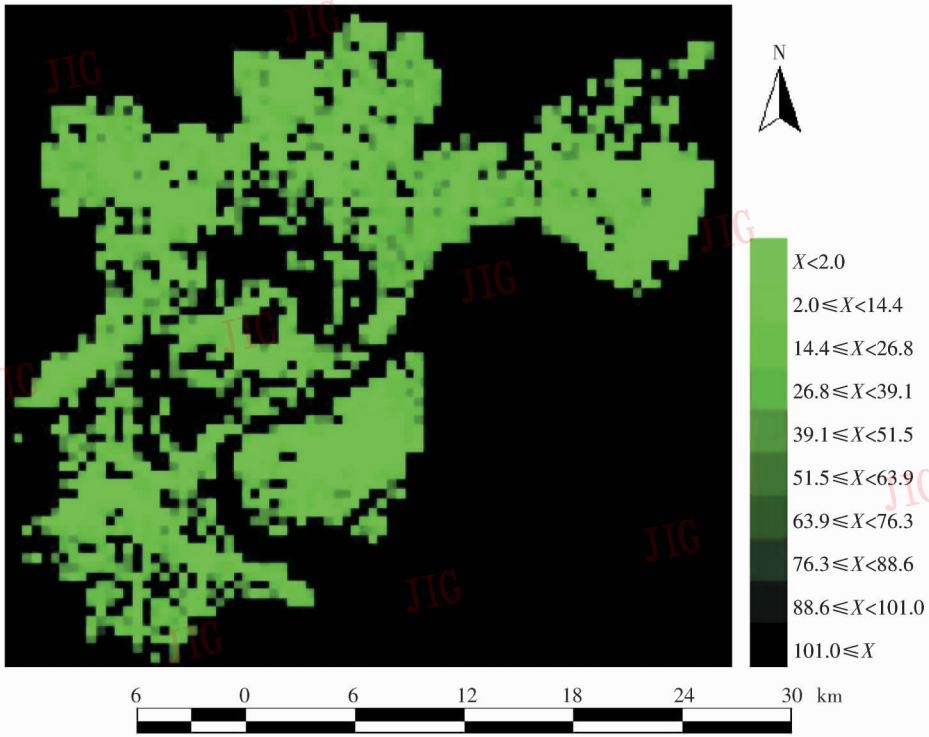


图 6 异步元胞演化间隔结果

Fig.6 Map of asynchronous cell

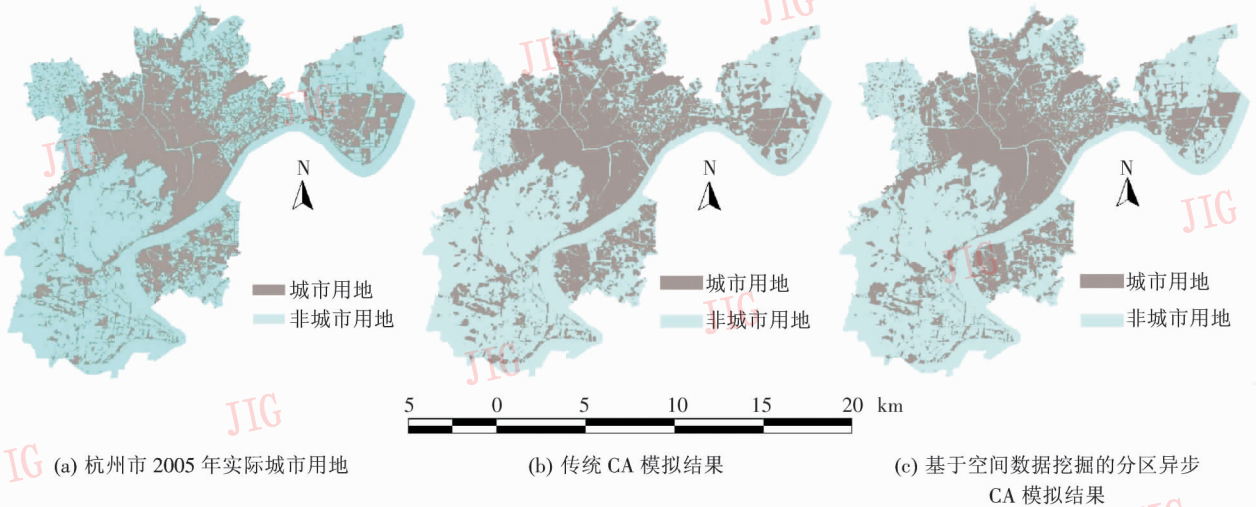


图 7 基于空间数据挖掘的分区异步 CA 模拟结果

Fig.7 Simulation results of synchronous and asynchronous CA

2.3 精度检验

为了验证元胞自动机模型模拟结果的正确性与可靠性,从两个方面对基于空间数据挖掘的分区异步 CA 的模拟结果进行了验证:一方面,采用逐点对比的方法对基于空间数据挖掘的分区异步 CA 模拟结果进行了评估;另一方面,通过度量空间对象集中和分散程度的 Moran I 指数来分析模拟结果在空间

整体结构和形态方面与杭州市 2005 年实际的土地利用情况的相似性。

2.3.1 逐点对比精度

逐点对比法是元胞自动机模型精度检验的最直观有效的方法,也是最常用的元胞自动机模型精度评估方法之一。它直接将元胞自动机模型的模拟结果与实际的土地利用情况进行叠加,就可以直观地

了解到元胞自动机模型的模拟精度。表1为传统CA模拟结果。

表1 传统CA模拟结果的混淆矩阵

Tab.1 Comparing matrix of traditional CA

		2005年模拟结果		精度/%
		城市用地 (元胞数)	非城市用地 (元胞数)	
2005年 实际情况	城市用地 (元胞数)	98 791	18 006	84.58
	非城市用地 (元胞数)	17 830	147 864	89.24
总精度				87.31

从表2可以看出,基于空间数据挖掘的分区异步元胞自动机模型的模拟结果与传统的元胞自动机模型相比,精度有明显的提高,城市用地的模拟精度高达87.27%,非城市用地的模拟精度高达91.04%,总精度也高达89.49%。因此,分区和异步可以很好地反映地理现象的空间差异性以及地理过程演化速率的空间差异性,所以基于空间数据挖掘的分区异步元胞自动机模型可以较大地提高元胞自动机模型进行土地利用模拟的精度。

表2 基于空间数据挖掘的分区异步CA模拟结果的混淆矩阵

Tab.2 Comparing matrix of partitioned & asynchronous CA result

		2005年模拟结果		精度/%
		城市用地 (元胞数)	非城市用地 (元胞数)	
2005年 实际情况	城市用地 (元胞数)	101 758	14 840	87.27
	非城市用地 (元胞数)	14 863	151 030	91.04
总精度				89.49

2.3.2 整体对比精度

逐点对比精度检验可以从一方面说明元胞自动机模型模拟结果与实际土地利用情况的接近程度,但却不能很好地反映模拟结果的空间形态和整体结构与实际土地利用情况的接近性。为了衡量模拟结果在空间形态和整体结构方面的精度,引入Moran I指数来对元胞自动机模拟结果的空间形态和结构进行评估。评估结果如表3所示。

从表3中可以看出,基于空间数据挖掘的分区异步CA模拟结果的Moran I指数比传统CA模拟结

表3 Moran I指数对比表

Tab.3 Comparing table of Moran I indices

	杭州市2005年 实际土地利用	基于统一转换规则 CA模拟结果	基于双约束空间聚类 分区异步CA模拟结果
Moran I	0.461	0.508	0.473

果更接近真实土地利用情况的Moran I指数,这说明基于空间数据挖掘的分区异步CA具有更好的模拟土地利用空间形态和总体结构的能力。

3 结论

地理现象的演化存在着空间差异性,一方面地理现象的演化规律存在着空间差异性;另一方面地理现象的演化速率存在着空间差异性。作为复杂地理现象模拟的有力手段,元胞自动机模型的模拟结果往往也会受到地理现象空间差异性的影响。然而,大多数元胞自动机模型在模拟地理现象时都采用统一的转换规则和相同的演化速率进行演化,因此,元胞自动机模型模拟结果的精度受到了限制。针对这一问题,提出了基于空间数据挖掘的分区异步元胞自动机模型,采用双约束空间聚类的方法对元胞空间进行划分,使同一分区内的元胞既在空间位置上相互邻近,又具有相对一致的非空间属性信息;通过标准格网划分方法求取每一个格网的元胞演化速率,使元胞自动机模型的演化更接近实际地理现象的变化。以杭州市土地利用变化为例对基于空间数据挖掘的分区异步元胞自动机模型进行了实证研究,结果表明:与传统的元胞自动机模型相比,基于空间数据挖掘的分区异步元胞自动机模型具有较高的模拟精度,并且适合用于较大区域较长时间段地理现象的动态变化模拟。

在基于空间数据挖掘的分区异步元胞自动机模型中,元胞空间的分区和元胞演化速率的确定是十分重要的。本文采用双约束聚类的方法对元胞空间进行划分,考虑了空间和非空间两方面的因素,聚类结果符合实际情况,较好地反映了地理现象的空间异质性原理。在采用双约束空间聚类进行元胞空间分区时,权重 w_p 和 w_a 的确定直接关系到分区结果的好坏,本文根据经验采用试算的方法确定权重 w_p 和 w_a 以保证较好的分区效果,今后需要对权重 w_p 和 w_a 的确定方法进行研究,提出科学合理的方法。在确定元胞演化速率时,采用了标准格网划分的方法获取速率格网,以后还需要更深入地研究。同时,

基于空间数据挖掘的分区异步元胞自动机模型在模拟较大区域较长时间的地理现象动态演变方面的应用有待于进一步的研究。

参考文献 (References)

- [1] Li Xia, Ye Jia'an, Liu Xiaoping, et al. Geographical Simulation Systems: Cellular Automata and Multi-Agent Systems [M]. Beijing: Science Press, 2007: 1. [黎夏, 叶嘉安, 刘小平, 等. 地理模拟系统: 元胞自动机与多智能体 [M]. 北京: 科学出版社, 2007: 1]
- [2] Liu Yaolin, Liu Yanfang, Ming Dongping. Urban expansion model based on cellular automata of decision-making with grey situation [J]. Geomatics and Information Science of Wuhan University, 2004, 29(1): 7-13. [刘耀林, 刘艳芳, 明冬萍. 基于灰色局势决策规则的元胞自动机城市扩展模型 [J]. 武汉大学学报: 信息科学版, 2004, 29(1): 7-13.]
- [3] Liu Xiaoping, Li Xia. Fisher discriminant and automatically getting transition rule of CA [J]. Acta Geodaetica et Cartographica Sinica, 2007, 36(1): 112-118. [刘小平, 黎夏. Fisher 判别及自动获取元胞自动机的转换规则 [J]. 测绘学报, 2007, 36(1): 112-118.]
- [4] Zhang Yan, Chen Yunhao, Li Jing. Study on the dynamic city model based on cellular automata and population condition [J]. Journal of Image and Graphics, 2007, 12(8): 1483-1488. [张岩, 陈云浩, 李京. 人口条件约束下的城市元胞自动机模型研究 [J]. 中国图象图形学报, 2007, 12(8): 1483-1488.]
- [5] Batty M, Xie Yichun, Sun Zhanli. Modeling urban dynamics through GIS-based cellular automata [J]. Computers, Environment and Urban Systems, 1999, 23(3): 205-233.
- [6] Almeida C M D, Batty M, Antonio M V M, et al. Stochastic cellular automata modeling of urban land use dynamics [J]. Computer, Environment and Urban Systems, 2003, 27(5): 481-509.
- [7] Sun Tengda, Wang Jinfeng. Microscopic traffic simulation model based on road network grid [J]. Journal of System Simulation, 2007, 19(12): 2735-2739. [孙腾达, 王劲峰. 一种基于路网网格化的微观交通仿真模型 [J]. 系统仿真学报, 2007, 19(12): 2735-2739.]
- [8] Huang Guangqiu, Wang Xiaohai, Chen Huiming. New method of computer simulation on fire spreading along underground tunnels system based on cellular automata [J]. Journal of System Simulation, 2007, 19(1): 201-205. [黄光球, 汪晓海, 陈惠明. 基于元胞自动机的地下矿火灾蔓延仿真方法 [J]. 系统仿真学报, 2007, 19(1): 201-205.]
- [9] Yang Qingsheng, Li Xia. Cellular automata for simulating land use changes based on support vector machine [J]. Journal of Remote Sensing, 2006, 10(6): 836-846. [杨青生, 黎夏. 基于支持向量机的元胞自动机及土地利用变化模拟 [J]. 遥感学报, 2006, 10(6): 836-846.]
- [10] Jenerette G D. Analysis and simulation of land-use change in the central Arizona-Phoenix region, USA [J]. Landscape Ecology, 2001, 16(7): 611-626.
- [11] Liu X H, Andersson C. Assessing the impact of temporal dynamics on land use change modeling [J]. Computer, Environment and Urban Systems, 2004, 28(1-2): 107-124.
- [12] Takeshi A, Tetsuya A. Empirical analysis for estimating land use transition potential functions-case in Tokyo metropolitan region [J]. Computer, Environment and Urban Systems, 2004, 28(1-2): 65-84.
- [13] Liu L, Stuart R P. Modeling urban development with cellular automata incorporating fuzzy-set approaches [J]. Computers, Environment and Urban Systems, 2003, 27(5): 637-658.
- [14] Claudia M A, Batty M. Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation [J]. Computers Environment and Urban Systems, 2003, 27(5): 481-509.
- [15] Li Xia, Ye Jia'an. Neural-network-based cellular automata for realistic and idealized urban simulation [J]. Acta Geographica Sinica, 2002, 57(2): 159-166. [黎夏, 叶嘉安. 基于神经网络的单元自动机 CA 及真实和优化的城市模拟 [J]. 地理学报, 2002, 57(2): 159-166.]
- [16] Liu X P, Li X, Yeh A G O. Discovery of transition rules for geographical cellular automata by using ant colony optimization [J]. Science in China: Series D, 2007, 50(10): 1578-1588.
- [17] Hu Maogui, Fu Xiaoyang, Zhang Shuqing, et al. Land-cover simulation and forecast of melmeg wetland using cellular automata [J]. Resource Science, 2007, 29(2): 142-148. [胡茂桂, 傅晓阳, 张树清, 等. 基于元胞自动机的莫莫格湿地土地覆被预测模拟 [J]. 资源科学, 2007, 29(2): 142-148.]
- [18] Yang Qingsheng, Li Xia. Integration of multi-agent systems with cellular automata for simulation urban land expansion [J]. Scientia Geographica Sinica. [杨青生, 黎夏. 多智能体与元胞自动机结合及城市用地扩张模拟 [J]. 地理科学, 2007, 27(4): 542-548.]
- [19] Zhou Jiaogen. Study on Dual-constraint and Alterable Size Spatial Cluster and Distributed Isolated Points Detection [D]. Wuhan: Wuhan University, 2007. [周脚根. 双约束和粒度可变空间聚类以及分布式离群点检测研究 [D]. 博士学位论文. 武汉: 武汉大学, 2007.]
- [20] Jiao Limin, Liu Yaolin, Ren Zhouqiao. Spatial points clustering based on self-organizing neural networks and its application [J]. Geomatics and Information Science of Wuhan University, 2008, 33(2): 168-171. [焦利民, 刘耀林, 任周桥. 基于自组织神经网络的空间点群聚类及其应用分析 [J]. 武汉大学学报: 信息科学版, 2008, 33(2): 168-171.]