

# 基于模糊粗糙集的 GIS 数据转换方法研究

廖伟华

(广西大学数学与信息科学学院, 广西 南宁 530004)

**摘要:**GIS 按照数据结构可以分为矢量数据和栅格数据两种。GIS 的空间建模大部分是基于栅格数据进行的, 因此, 对于一些在矢量结构中表达的属性(如人口密度等), 需要转换成栅格数据。传统的数据转换方法一般都是面积占优法, 这种方法没有充分考虑空间数据的模糊性问题。采用模糊粗糙集的理论, 对矢量多边形数据转换成栅格数据的方法做了探讨。基于模糊粗糙集的数据转换方法考虑到了多边形数据的边界模糊性, 利用隶属函数来确定多边形边界的不确定性。

**关 键 词:**模糊粗糙集; GIS; 数据转换

**中图分类号:**TP 274 **文献标识码:**A **文章编号:**1004-0323(2007)06-0759-03

## 1 引 言

GIS 空间数据结构分为矢量数据结构和栅格数据结构两种, 这两种数据结构各有优缺点。矢量数据具有面向对象的数据表示、数据结构紧凑、冗余度低等特点。栅格数据数据结构简单、有利于空间分析和地理现象的模拟<sup>[1]</sup>。因此, 在 GIS 中进行空间数据建模时一般都采用的是栅格数据, 而面对一些具有矢量特性的数据, 如土壤类型、人口密度等。则需要转换成栅格数据进行模拟计算。

目前矢量数据能转换成栅格数据的基本元素分为点、线、面。对于点和线由于都是没有宽度的两种元素, 在点和线转换成栅格时, 对所经过的栅格转换成相应的属性值的栅格单元。没有经过的栅格转换成空值。而对于多边形, 由于多边形具有面积属性, 在多边形转换成栅格数据时, 在多边形内部会转换成具有相应属性值的栅格单元。但在多边形的边界, 由于边界不具有面积, 在边界转换成栅格数据时, 具有模糊性<sup>[2~4]</sup>。目前 GIS 中矢量多边形转换成栅格数据有颜色填充法、掩膜法、边界代数多边形填充法、边界追踪法以及折线边界追踪法等。这些算法都是从如何快速实现矢栅转换的角度所做的工作, 没有考虑到临界栅格点的取值方法对多边形栅格化结果的影响。在临界栅格点的取值方法上, 基本上采用栅格像元“面积占优”的原则, 即对边界线

穿过的这部分栅格点, 取面积占多数的多边形的值为该处栅格的值<sup>[5]</sup>。而面积误差最小约束考虑了边界的模糊性, 但是对于边界栅格的隶属没有一个很好的控制办法。

## 2 基本理论

由于地理现象和过程的复杂性, 很难准确区别这些现象的边界。如对于多边形, 在转换成栅格数据时, 很难判断它的边界在转换时是否转换成与多边形具有相同属性的栅格数据。

### 2.1 粗糙集与模糊集比较

模糊集和粗糙集在处理不确定问题时推广了经典集合理论, 它们都可以用来描述知识的不确定性和不完全性。然而它们的侧重点不同。

从知识的“粒度”上来看, 模糊集主要着眼于知识的模糊性, 而粗糙集着眼于知识的粗糙性。从知识的描述方法上来看, 模糊集是通过对象关于集合的隶属程度来近似描述的, 而粗糙集是通过一个集合关于某个已知的可利用的信息库的一对一、下近似来描述的。从集合对象间的关系来看, 模糊集强调的是集合边界的变态定义, 即边界的不分明性; 而粗糙集强调的是集合对象间的不可分辨性。从研究的对象来看, 模糊集研究的是属于同一类的不同对象的隶属关系, 重在隶属度; 而粗糙集研究的是不同类中的对象组成的集合之间的关系, 重在分类<sup>[6~8]</sup>。

因此,这两种理论有很强的互补性,将这两种理论进行局部整合后,能很好地处理一些具有不确定性和不完全性的知识。

2.2 粗糙模糊集理论

模糊粗糙集是在论域  $U$  中,用  $A$  关于  $(U,R)$  的一对上下近似来描述在经典粗糙集中不能精确描述的经典集合  $A$ 。

定义 1 给定知识库  $K=(U,R)$ ,对于每个子集  $X\subseteq U$  和一个等价关系  $R\in \text{ind}(K)$ ,定义两个子集:

$$\begin{aligned}RX &= \bigcup \{Y \in U \mid /R \mid Y \subseteq X\} \\ \overline{R}X &= \bigcup \{Y \in U \mid /R \mid Y \cap X \neq \emptyset\}\end{aligned}\tag{1}$$

分别称它们为  $X$  的  $R$  下近似集和  $R$  上近似集。

定义 2 设  $(U,R)$  是 Pawlak 近似空间,即论域  $U$  上的一个等价关系。若  $A$  是  $U$  上的一个模糊集合,则  $A$  关于  $(U,R)$  的一对下近似  $\underline{A}_R$  和上近似  $\overline{A}_R$  定义为  $U$  上的一对模糊集合,其隶属函数分别定义为:

$$\begin{aligned}\underline{A}_R(x) &= \inf\{A(y) \mid y \in [x]_R\}, x \in U \\ \overline{A}_R(x) &= \sup\{A(y) \mid y \in [x]_R\}, x \in U\end{aligned}\tag{2}$$

其中: $[x]_R$  为元素  $x$  在关系  $R$  下的等价类。若  $\underline{A}_R = \overline{A}_R$ ,则称  $A$  是可定义的,否则称  $A$  是模糊粗糙集。称  $\underline{A}_R$  是关于  $(U,R)$  的正域,称  $\sim \overline{A}_R$  是  $A$  关于  $(U,R)$  的负域,称  $\overline{A}_R \cap (\sim \underline{A}_R)$  为  $A$  的边界。

定义 3 设  $(U,R)$  是 Pawlak 近似空间, $A$  是  $U$  上的模糊集合,则  $A$  关于近似空间  $(U,R)$  依参数  $0<\beta\leq\alpha\leq 1$  的下近似  $\underline{A}_\alpha$  和上近似  $\overline{A}_\beta$  分别定义为:

$$\begin{aligned}\underline{A}_\alpha &= \{x \in U \mid \underline{A}(x) \geq \alpha\} \\ \overline{A}_\beta &= \{x \in U \mid \overline{A}(x) \geq \beta\}\end{aligned}\tag{3}$$

$\underline{A}_\alpha$  可以解释为  $U$  中肯定属于模糊集  $A$  的隶属度不小于  $\alpha$  的那些对象的全体, $\overline{A}_\beta$  可以解释为  $U$  中可能属于模糊集  $A$  的隶属度不小于  $\beta$  的那些对象的全体。

定义 4 设  $(U,R)$  是 Pawlak 近似空间, $A$  是  $U$  上的模糊集合,对于  $0<\beta\leq\alpha\leq 1$ ,定义  $A$  在近似空间  $(U,R)$  中关于参数  $\alpha,\beta$  的粗糙度如下:

$$\rho_{\alpha,\beta}^A = 1 - \frac{|\underline{A}_\alpha|}{|\overline{A}_\beta|}\tag{4}$$

约定当  $\overline{A}_\beta = \phi$  时,  $\rho_{\alpha,\beta}^A = 0$ 。

3 实例分析

设有如图 1 所示多边形,将图中多边形转换成栅格数据时,按照系统软件默认的面积占优法,按照一个单位的栅格单元大小,得到的栅格数据共 44 个栅格。此方法只不过是模糊集的某一  $\lambda$  截集,如常

规方法算出的结果就是  $\lambda = 0.5$  时的结果<sup>[9]</sup>。

用粗糙集来转换多边形:下近似集为 35 个栅格单元,上近似集为 63 个栅格单元。也就是说按照粗糙集理论,上述多边形在进行数据转换时将有 28 个属于边界域的模糊栅格。这 28 个可能可以转换成属于多边形的栅格数据,也可以转换成不属于多边形的栅格数据。其中粗糙度为 0.444。

图 1 GIS 矢量地图

Fig. 1 GIS vector map

用模糊粗糙集来进行数据转化时,先要确定隶属函数。隶属函数的确定由边界转换时,模糊边界所占栅格单元的面积比例确定。

按照模糊粗糙集理论,需要先确定  $\alpha,\beta$ ,本试验分别对这两个参数取 0.6、0.3 和 0.7、0.4 两次进行试验(在实际操作中可以根据需要自己选取参数),对多边形进行数据转换。当参数取 0.6、0.3 时进行数据转换所得结果如图 2 所示,其中  $|\underline{A}_\alpha|$  为 41,  $|\overline{A}_\beta|$  为 56,其中关于参数  $\alpha,\beta$  的粗糙度为 0.2679。当参数取 0.7、0.4 时  $|\underline{A}_\alpha|$  为 41,  $|\overline{A}_\beta|$  为 49,其中关于参数  $\alpha,\beta$  的粗糙度为 0.1633,得到结果如图 3 所示。

图 2 模糊粗糙集转换结果( $\alpha=0.6,\beta=0.3$ )

Fig. 2 Result of fuzzy rough set ( $\alpha=0.6,\beta=0.3$ )

合 GIS 空间数据模糊性的特性。因此,基于模糊粗糙集的转换方法,通过隶属函数既可以考虑到面积占优法,也可以考虑到面积误差最小约束,还充分考虑到了空间数据模糊性,转换精度可以根据实际需要完全可控。

参考文献:

[1] 黄杏元,马劲松,汤勤. 地理信息系统概论[M]. 北京:高等教育出版社,2002.

[2] 杜世宏,王桥,李顺,等. 模糊对象粗糙表达及其空间关系研究[J]. 遥感学报,2004,8(1):1-8.

[3] 李三平,葛咏,李德玉,等. 遥感信息不确定性度量[J]. 山西大学学报(自然科学版),2006,29(1):30-32.

[4] 朱德海,严泰来,杨永侠. 土地管理信息系统[M]. 北京:中国农业大学出版社,2002.

[5] 王晓理,孙庆辉,江成顺. 面积误差最小约束下矢量数据向栅格数据转换的优化算法[J]. 测绘学报. 2006. 35(3):273-277.

[6] 张文修,吴伟志,梁吉业,等. 粗糙集理论与方法[M]. 北京:科学出版社,2005.

[7] 黄正华,胡宝清. 模糊粗糙集理论研究进[J]. 模糊系统与数学, 2005,19(4):125-134.

[8] 贺仲雄. 模糊数学及其应用[M]. 天津:天津科学技术出版社, 1983.

[9] 胡圣武,王新洲,潘正风,等. GIS 中的模糊不确定性以及处理方法[J]. 武汉大学学报(信息科学版),2005,30(5):417-459.

图 3 模糊粗糙集转换结果( $\alpha=0.7, \beta=0.4$ )  
Fig.3 Result of fuzzy rough set ( $\alpha=0.7, \beta=0.4$ )

4 结 语

GIS 软件在矢量数据转换成栅格数据时,对于多边形转换按照面积占优方法来处理。所以没有充分考虑 GIS 中多边形的边界模糊性。而如果按照粗糙集传统上近似,下近似转换,转换结果也只有两种,一定程度上也没能考虑边界的模糊性。而基于模糊粗糙集的理论进行数据格式转换,参照隶属函数。用隶属函数来确定多边形边界的不确定性。不同的隶属函数可以得到不同的结果。分析过程符

Method Study of GIS Data Transformation Based on Fuzzy Rough Set

LIAO Wei-hua

(Department of Mathematics and Infomation ,Guangxi University, Nanning 530004,China)

**Abstract:** The data of GIS are divided into vector data and grid data for its structure. The spatial model of GIS is mainly based on grid data. Therefore, some vector data with special properties should be transformed to grid data and area occupied superiorly method is widely used. However, this method did not fully consider the fuzzy attribute of spatial data. Rough set theory is used to discuss the potential approach to convert polygon data to grid data, meanwhile, using the membership function and fuzzy rough set to determine the uncertainty for the boundary of polygon and to transform the polygon to grid.

**Key words:** Fuzzy rough set, GIS, Data transformation