

文章编号:1671-8860(2009)03-0362-04

文献标志码:A

# 基于信息损失的街道渐进式选取方法

田晶<sup>1,2</sup> 郭庆胜<sup>1,2</sup> 冯科<sup>1,2</sup> 马盟<sup>1,2</sup>

(1 武汉大学资源与环境科学学院,武汉市珞喻路129号,430079)

(2 武汉大学地理信息系统教育部重点实验室,武汉市珞喻路129号,430079)

**摘要:**对街道网的量测信息、拓扑信息和专题信息进行了计算,在此基础上提出了一种基于信息损失的街道选取方法。按照该方法对GB 12343-90中的一幅街区式居民地地图进行了实验。结果表明,该方法能够动态和渐进地进行街道选取。

**关键词:**街道选取;优化;信息损失;渐进

**中图法分类号:**P283.7; P208

街道是居民地内的各种通道,街道选取是居民地综合的一个重要步骤。选取街道需要综合考虑街道的语义、几何和拓扑三个方面的属性。渐进式地图综合强调在细节减少的过程中实时地对每一个步骤进行评价,使得每一步所得的结果在当前的抽象层次上是可用的,是进一步综合的基础<sup>[1]</sup>,而如何进行评价是整个问题的关键。Li<sup>[2]</sup>指出,地图综合的研究将从要素级别转向类级别和地图级别。在地图级别上,地图信息的度量以及综合对地图信息量的影响是关键问题。本文提出了一种基于信息损失的渐进式街道选取方法。

## 1 街道网信息的量化与信息损失

在地图信息论的发展过程中,对地图信息的理解有两种代表性观点:①信息是信源与接受信息者之间的交互;②信息是地图的直接属性。本文所提到的信息是指后一种观点,即街道网信息是街道网的直接属性。从20世纪60年代起,许多学者对地图信息的量化进行了研究<sup>[3-8]</sup>。从这些研究中可知,对地图信息量化的关键是如何定义地图目标间的差别。

假定对于街道网有两种要素类型(即街道和街区)和三种专题类型(即主要街道、次要街道和街区),街道网由X条街道(X<sub>1</sub>条主要街道,X<sub>2</sub>条次要街道)和Y个街区组成。本文对街道网的量

测信息、拓扑信息和专题信息进行了计算。图1是某街道网中轴线图。该街道网有8个街区(A,B,C,D,E,F,G,H)和11条街道,其中,有4条主要街道(1,2,3,4)、7条次要街道(5,6,7,8,9,10,11)。

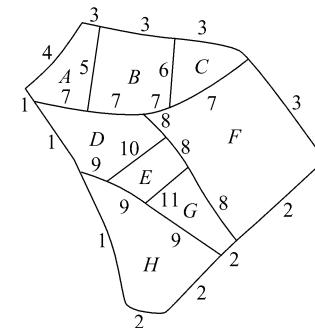


图1 街道网示意图

Fig. 1 A Partial Map of Street Network

### 1.1 量测信息

定义1 街道网的量测信息主要考虑街道间的长度差别和街区间的面积差别。

其量化方法如下:选择到第*i*条街道的概率为 $P_i = l_i / L, i=1, 2, \dots, X$ ,其中, $l_i$ 是第*i*条街道的长度, $L$ 是街道长度的总和。街道的量测信息为:

$$I(\text{SMetric}) = - \sum_{i=1}^X P_i \log_2 P_i$$

对于街区的量测信息,主要考虑街区面积占街

区总面积的比例,选择到第  $j$  个街区的概率为  $P_j = a_j/A, j=1,2,\dots,Y$ , 其中,  $a_j$  为第  $j$  个街区的面积,  $A$  是街区面积的总和。街区的量测信息为:

$$I(\text{BMetric}) = -Y \sum_{j=1}^Y P_j \log_2 P_j$$

街道网的量测信息为:

$$H(\text{Metric}) = H(\text{SMetric}) + H(\text{BMetric})$$

## 1.2 拓扑信息

定义 2 街道网的拓扑信息主要考虑街道间的连通性差别和街区间邻近性差别。

其量化方法如下:沿用 Neumann<sup>[5]</sup>提出的拓扑信息量化方法,街道网拓扑信息的计算依赖于两种对偶图,对图 1 所示街道网的街道对偶图和街区对偶图如图 2 所示。

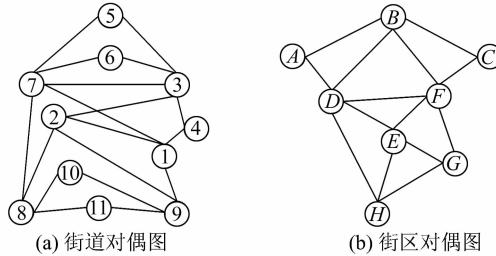


图 2 街道对偶图和街区对偶图

Fig. 2 Dual Graph of Streets and Dual Graph of Blocks

街道对偶图含有  $X$  个结点,结点代表街道,边代表街道相交。对结点进行分类,分类规则为:结点的度相等;与结点相邻的结点的度对应相等。考虑图 2(a)中的结点④、⑤和⑩,它们的度均等于 2。与结点相邻的是一个度等于 4 的结点(①)和一个度等于 5 的结点(③);与结点⑤相邻的结点是两个度等于 5 的结点(③和⑦);与结点⑩相邻的结点是两个度等于 4 的结点(⑧和⑨)。所以,结点④、⑤和⑩属于不同的类别。假定第  $i$  类结点的数量为  $x_i$ ,则选择到第  $i$  类结点的概率为  $P_i = x_i/X$ ,街道拓扑信息为:

$$H(\text{STopo}) = -X \sum_{i=1}^k P_i \log_2 P_i$$

街区对偶图含有  $Y$  个结点,结点代表街区,边代表街区相邻,分类规则为:结点的度相等;与结点相邻结点的度对应相等。假定第  $j$  类结点的数量为  $y_j$ ,则选择到第  $j$  类结点的概率为  $P_j = y_j/Y$ ,街区拓扑信息为:

$$H(\text{BTopo}) = -Y \sum_{i=1}^k P_i \log_2 P_i$$

街道网的拓扑信息为:

$$H(\text{Topological}) = H(\text{STopo}) + H(\text{BTopo})$$

## 1.3 专题信息

定义 3 街道网的专题信息主要考虑目标间的邻居个数和类型的差别。

其量化方法如下:沿用 Li 等<sup>[2]</sup>提出的专题信息量化方法,他考虑的是基于 Voronoi 图的邻居。本文判定目标的邻居的方法是:① 街道与街道是邻居,则街道与街道相交;② 街道与街区是邻居,则街道是街区多边形的组成部分;③ 街区与街区是邻居,则街区与街区存在公共街道。邻居的专题类型分为三类:主要街道、次要街道和街区。假定第  $i$  个地图要素有  $N_i$  个邻居和 3 种邻居类型,第  $j$  类邻居有  $N_{ij}$  个,则第  $i$  个地图要素的邻居属于第  $j$  类的概率是  $P_j = N_{ij}/N_i (j=1,2,3)$ ,第  $i$  个地图要素的专题信息为:

$$H_i(\text{Thematic}) = H(P_1, P_2, P_3) = -\sum_{j=1}^3 P_j \log_2 P_j$$

对于含有  $X+Y$  个地图要素的专题信息为:

$$H(\text{Thematic}) = \sum_{i=1}^{X+Y} H_i(\text{Thematic})$$

整个街道网的信息量为:

$$\text{Information} = H(\text{Metric}) + H(\text{Topological}) + H(\text{Thematic})$$

按照上述信息量化方法对图 1 所示的街道网的信息量进行计算,其量测信息为 57.5375 bit,拓扑信息为 52.0537 bit,专题信息为 26.1419 bit,信息总量为 135.7331 bit。

## 1.4 信息损失

地图综合是对地图信息的压缩,不可避免地涉及到信息的损失<sup>[9]</sup>。被删除街道所含有的信息消失了,街道网结构被简化,简化后的街道网信息量减少了。信息损失定义为初始状态下的信息量与当前状态下信息量的差值。删除图 1 中的某一条街道所导致的信息损失如表 1 所示。

## 2 街道选取的渐进式方法

一些学者将地图综合的过程看作优化过程<sup>[10-13]</sup>。地图综合作为优化过程的关键是搜索空间和评价函数的确定。  
 ① 搜索空间由各种可能的实现构成。针对街道选取问题,本文的搜索空间由街道网的所有街道构成,每次删除街道后,搜索空间被更新,由现有状态街道网中的街道构成。  
 ② 评价函数对搜索空间中的元素的相对质量进行评价,评价函数保证每次删除街道后的信息损失最小。  
 ③ 核心思想是,对街道进行试探性删除,计算各种删除状态下的信息损失,删除对应于

表 1 信息损失

Tab. 1 Information Loss

删除的街道 ID	量测信息	拓扑信息	专题信息	信息总量	信息损失
1	0	0	0	0	135.733 1
2	0	0	0	0	135.733 1
3	0	0	0	0	135.733 1
4	0	0	0	0	135.733 1
5	48.959 8	36.115 9	23.312 7	108.388 4	27.344 7
6	49.147	36.115 9	23.278 6	108.541 5	27.191 6
7	12.470 4	7.609 6	8.632 1	28.712 1	107.021 0
8	31.762 5	28.854 7	18.585 3	79.202 5	56.530 6
9	32.390 0	23.609 6	19.075 7	75.075 3	60.657 8
10	49.294 9	50.870 8	24.281 8	124.447 5	11.285 6
11	49.666 3	50.870 8	24.281 8	124.818 9	10.914 2

信息损失最小的街道,对街道网更新。重复该过程,直到满足综合目标,如街区面积大于给定阈值、保留一定百分比的街道长度等,它们都是尺度对应的函数。本文采用保留一定百分比的街道长度作为过程的终止条件。

如表 1 所示,信息损失最小的是 11,所以第一步删除街道 11,删除街道后对应的状态成为下一次删除的初始状态。重复该过程,直到满足综合目标。删除图 1 中次要街道的顺序如表 2 所示。

表 2 街道渐进式选取结果

Tab. 2 Result of Progressive Selection of Streets

步骤	删除的街道 ID	量测信息	拓扑信息	专题信息	信息总量	信息损失
1	11	49.666 3	50.870 8	24.281 8	124.818 9	10.914 2
2	10	41.354 6	36.039 1	21.619 0	99.012 7	25.806 2
3	8	31.762 5	28.854 7	18.585 3	79.202 5	19.810 2
4	6	25.353 1	19.651 5	15.652 5	60.657 1	18.545 4
5	5	19.288 0	13.509 8	11.946 4	44.744 2	15.912 9
6	7	12.470 4	7.609 6	8.632 1	28.712 1	16.032 1
7	9	7.555 4	0	3.673 2	11.228 6	17.483 5

实验在 Microsoft Visual Studio 开发环境下用 C++ 语言结合 ESRI 的组件产品 ArcObjects 进行。实验数据来源于 GB 12343-90<sup>[14]</sup> 中的 1 : 25 000 城市居民地地图的一部分。图 3 显示了实验的部分结果。

计算是将三种信息量进行简单相加,三种信息量之间是否存在一定的权重关系没有得到证明;② 算法的实质是贪心搜索,可能会导致局部优化问题;③ 综合算法与尺度的关系是隐式对应的,综合后的实际尺度不可知。

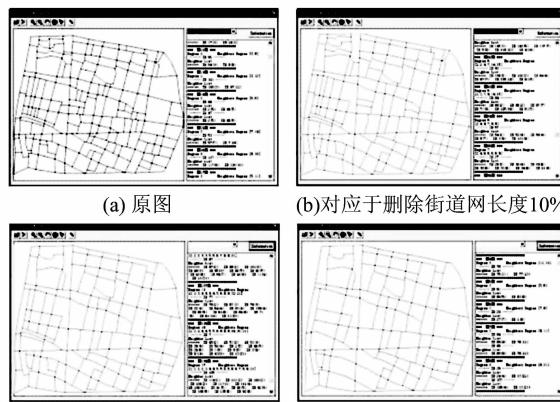


图 3 街道网原图及不同阈值的综合结果

Fig. 3 Street Network and Its Three Generalized Versions

该算法存在如下一些问题:① 整体信息量的

### 3 结语

本文所提出的街道选取方法具有渐进性和动态性,渐进性表现在通过逐条删除街道来实现最终的选取目标;动态性表现在每次删除街道后,街道网对应新的状态,需要重新计算信息量,搜索空间发生变化。本文应用信息论解决如何进行综合的问题,所提出的算法是地图级别的综合算法。进一步的研究工作主要有:① 建立尺度与算法输入参数间的显式对应关系;② 街道网矢量数据的渐进式传输。

### 参 考 文 献

- [1] 郭庆胜,黄远林,郑春燕,等.空间推理与渐进式地图综合[M].武汉:武汉大学出版社,2007

- [2] Li Zhilin. Digital Map Generalization at the Age of Enlightenment: a Review of the first Forty Years [J]. *The Cartographic Journal*, 2007, 44(1): 80-93
- [3] Sukhov V I. Information Capacity of a Map Entropy [J]. *Geodesy and Areophotography*, 1967(5): 212-215
- [4] Sukhov V I. Application of Information Theory in Generalization of Map Contents [J]. *International Yearbook of Cartography*, 1970(5): 41-47
- [5] Neumann J. The Topological Information Content of a Map: an Attempt at a Rehabilitation of Information Theory in Cartography [J]. *Cartographica*, 1994, 31(1): 26-34
- [6] Bjørke J T. Framework for Entropy-based Map Evaluation [J]. *Cartography and Geographical Information Systems*, 1996, 23(2): 78-95
- [7] Li Zhilin, Huang Peizhi. Quantitative Measures for Spatial Information of Maps [J]. *International Journals of Geographical Information Science*, 2002, 16(7): 699-709
- [8] Battersby S E, Clarke K C. Quantifying Information in Line Generalization [C]. The 21st International Cartographic Conference, Durban, South Africa, 2003
- [9] 郭庆胜. 地图自动综合理论与方法 [M]. 北京: 测绘出版社, 2002
- [10] Ware M, Jones C. Conflict Reduction in Map Generalization Using Iterative Improvement [J]. *GeoInformatica*, 1998, 2(4): 383-407
- [11] Harrie L E. The Constraint Method for Solving Spatial Conflicts in Cartographic Generalization [J]. *Cartography and Geographic Information Science*, 1999, 26(1): 55-69
- [12] Sester M. Optimization Approaches for Generalization and Data Abstraction [J]. *International Journals of Geographical Information Science*, 2005, 19(8/9): 871-897
- [13] Ware M, Wilson I, Ware J, et al. A Tabu Search Approach to Automated Map Generalization [C]. The 10th ACM International Symposium on Advances in Geographic Information Systems, McLean, Virginia, USA, 2002
- [14] 国家测绘局标准化研究所. 测绘标准汇编——地图制图及印刷卷 [M]. 北京: 中国标准出版社, 2003

**第一作者简介:**田晶,博士生。主要研究方向为地图自动综合、地理信息智能化处理。

E-mail:yutaka-2010@163.com

## Progressive Selection Approach of Streets Based on Information Loss

TIAN Jing<sup>1,2</sup> GUO Qingsheng<sup>1,2</sup> FENG Ke<sup>1,2</sup> MA Meng<sup>1,2</sup>

(1) School of Resources and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2) Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University,  
129 Luoyu Road, Wuhan 430079, China)

**Abstract:** The characteristics of existing approaches to the selection of streets are summarized based on an analysis of the literatures. The metric information, topological information and thematic information are calculated for street networks. From an artificial intelligence perspective, the generalization process can be seen as an optimization problem. A progressive approach to the selection of streets base on information loss is proposed. This approach minimizes the information loss. An experiment was conducted on a partial map of an urban settlement from the compilation specifications for 1 : 25 000, 1 : 50 000 topographic maps. The Results show that our proposed approach can select the streets progressively and dynamically.

**Key words:** selection of streets; optimization; information loss; progressive

**About the first author:** TIAN Jing, Ph.D candidate, majors in automated map generalization and intelligent handling of geographic information.

E-mail: yutaka-2010@163.com