

文章编号:1001-1595(2007)03-0351-07

中图分类号:P283.1

文献标识码:A

基于网眼密度的道路选取方法

胡云岗^{1,2}, 陈军², 李志林³, 赵仁亮²

(1. 中国矿业大学(北京校区)资源与安全工程学院, 北京100083; 2. 国家基础地理信息中心, 北京100044; 3. 香港理工大学土地测量与地理资讯学系, 香港九龙)

Selective Omission of Road Features Based on Mesh Density for Digital Map Generalization

HU Yun-gang^{1,2}, CHEN Jun², LI Zhi-lin³, ZHAO Ren-liang²

(1. Dept of Surveying and Land Science, China University of Mining and Technology, Beijing 100083, China; 2. National Geomatics Center of China, Beijing 100044, China; 3. Dept of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong, China)

Abstract: Selection of roads is an intractability generalization operation, due to the difficulty in retaining the density difference and connectivity of a road network. This paper proposes a new approach of selective omission for roads based on meshes density. Density of roads is calculated using meshes as units and represents local region density in a network. Usually, the density is different between maps with different scales. The density thresholds are acquired by theoretical analysis and empirical study for mesh density on different scale maps. For meshes with density over threshold, the largest density meshes are taken out by turns and one road segment of the mesh is eliminated. Meanwhile the taken mesh is merged with the adjacent mesh at a time. Road segment to be eliminated is the most important in all segments on the boundary of this mesh according to the parameters reflecting the importance of road segments. In this study, the meshes are classed depending on the types of road segment. For the different type meshes, their thresholds are different, which can preserve the density difference. The process of eliminating road segments and merging meshes would ensure the road network connectivity. This proposed approach considers topological, geometric and semantic properties of the road network. A road network is selective eliminated employing this approach, and the results of selection achieve better effects.

Key words: map generalization; selection omission; road networks; network meshes; mesh density

摘要: 道路数据中的网眼密度能反映局部区域的道路密集程度, 通过确定目标尺度要求的密度阈值, 比例尺缩小后能够标识出数据中需要取舍路段的网眼; 循环剥离密度最大的网眼, 利用反映路段重要性的参数及其优先级, 渐进筛选出舍弃的路段, 并完成与邻接网眼的合并; 得到的选取结果保持了道路网在密度、拓扑、几何及语义方面的重要特征, 从而提出一种新的道路选取方法。最后进行实验, 验证该方法的有效性。

关键词: 地图综合; 选取; 道路网; 道路网眼; 网眼密度

1 引言

道路选取是地图综合中的重要操作之一, 指在比例尺缩小后将数据中不符合目标尺度要求的冗余目标舍弃, 基本要求是根据大比例尺道路的

密度和等级、长度等语义、几何特征, 筛选出重要道路, 舍弃次要道路, 并保持选取后道路网的拓扑连通^[1-5]。

以往道路选取主要是基于方根模型^[1]和图论^[2-5]来实现。其方法是, 根据综合前后的比例尺, 由方根模型确定要素的选取率; 根据图论拓扑

收稿日期: 2006-09-14; 修回日期: 2007-04-06

基金项目: 国家自然科学基金国家杰出青年基金类港澳青年学者合作研究基金项目(40329001); 国家自然科学基金重点项目(40337055); 国家自然科学基金项目(40301042)

作者简介: 胡云岗(1975-), 男, 山西临汾人, 博士生, 主要从事地图综合及地图数据更新方面的研究。E-mail: hyg@cumt.edu.cn

连通性度量及道路长度等参数,确定要选取的具体目标,以保持道路网的连通性和保留重要道路。方根模型是经大量统计而得出的一般性规律,仅给制图员提供了选取时的指导性原则。在自动化选取时,如要给出确切的选取率,方根模型还不尽合理^[6,7]。

作者经过统计发现,选取与道路密度相关。一方面不同尺度上的道路密度不同;另一方面在密集区域舍弃的路段较多,而在稀疏区域舍弃的路段较少。这表明密度可以作为阈值来控制道路的选取。因而,根据道路网的疏密划分出一系列小区域,通过确定密度阈值,利用几何、语义等属性,就可以对区域内的路段进行取舍。本文试图基于道路密度来完成对道路的选取操作。

2 道路密度——从线密度到网眼密度

2.1 传统道路密度

道路密度一般是指一定区域内道路长度与该区域面积的比值,可理解为一种线密度,数学表达为

$$D = L / A \quad (1)$$

其中, D 表示道路密度, A 表示区域面积, L 表示区域内道路的总长度。

人为划分区域计算密度比较困难且不方便选取操作,为此作者引出了网眼密度的概念。

2.2 网眼密度的概念与计算

为便于讨论,将道路网纵横交错形成的最小闭合区块称为“道路网眼”,如图1(a)中的阴影部分标示了3个网眼。以网眼为单位计算的密度称为“网眼密度”,指包含网眼的最小区域内道路总长度与区域面积的比值。设有包含网眼的任意小的区域(见图1(c)),根据式(1)网眼密度表达为

$$D = \frac{P + L + \varepsilon}{A + A} \quad (2)$$

式中, P 是网眼边界上路段总长度; L 是网眼内悬垂路段的总长度; ε 为区域内的路段总长度; A 为网眼的面积; A 为区域的面积。

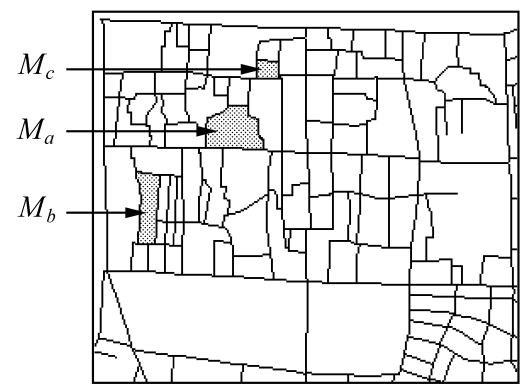
当 $\varepsilon \rightarrow 0$ 时,因为 $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{A + A} = 0$; 故

$$D_{lim} = \lim_{\varepsilon \rightarrow 0} \frac{P + L + \varepsilon}{A + A} = \frac{P + L}{A} \quad (3)$$

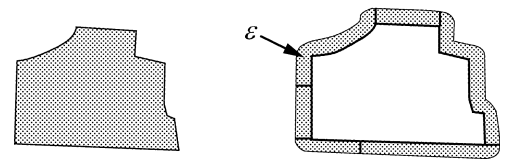
道路网络中的路段分为网眼边界路段和悬垂路段。实际上,式(3)中的网眼密度可简化为式(4):

$$D = P / A \quad (4)$$

即在计算网眼密度时,网眼内部的悬垂路段可以被忽略。其原因是:悬垂路段选取时不考虑道路网的连通性,仅依据几何与语义属性进行取舍,可以首先取舍悬垂路段;如果需要舍弃一个网眼上的边界路段,那么该网眼内的所有悬垂路段可能已经被舍弃了,这时考虑的仍然是网眼边界上的路段,因而简化是合理的。根据式(4),网眼 M_a 密度值为 0.0235 m/m^2 。



(a) 道路中的网眼



(b) 网上 M_a (c) 网眼 M_a 的领域

图1 网眼及其密度计算

Fig.1 Mesh and mesh density computing

3 基于网眼密度选取道路的基本原理

1. 网眼作为选取单元和网眼密度作为选取基准

网眼作为道路网自然划分的最小区域,形成了选取单元。当网眼密度大于阈值时,删除网眼上的路段,形成与邻接网眼的合并,新网眼密度也就变小。因而对这样的选取单元,密度阈值不是作为选取率使用,而是衡量路段是否密集的标准。

2. 保持道路网中不同区域的密度差别

以网眼边界上路段的最低等级为标准将道路网网眼分成不同类型。不同类型的网眼设置不同的密度阈值,有助于保持不同区域的密度差别。

3. 保持道路网的拓扑连通性

网眼作为一个闭合回路,舍弃一条或多条路段,并形成与邻接网眼的合并,不破坏道路网的连通。至于每次舍弃那些路段,需要根据路段的几何、语义等方面的参数,比较路段的重要性进行选择。对于悬垂路段的选取比较简单,仅需考虑路

段的几何与语义参数。

4 道路网眼密度阈值的确定

4.1 基于理论分析确定网眼密度阈值

网眼密度可以按实地尺寸计算,也可按图上尺寸计算。前者大小不随比例尺的变化而变化,以上网眼密度是以实地尺寸来计算的;而后者即网眼的图上密度,将随比例尺的缩小而变大,由 D_m 来表示。由于人眼分辨能力的局限性,网眼图上密度不能过大。对某一尺度而言,人眼刚好能分辨的最小单位网眼,其密度称为该尺度下允许的最大网眼密度(PLD),以 $D_{m \max}$ 来表示,实际意义就是最小的可视区域中能容纳下的最长道路。这样的网眼等同于Li和Openshaw提出的“自然法则”中的最小可视单元(Smallest Visible Object, SVO)。SVO可以用圆或栅格表示,用于简化单个地物要素或其组成部分^[10,11],则 $D_{m \max}$ 可表达如下

$$D_{m \max} = \frac{L_m / (L_m/2)^2}{4 L_m / (L_m)^2} \text{ 或 } = 4 / L_m \quad (5)$$

这里, L_m 表示SVO直径或边长的图上长度。

本文计算的是实地尺寸密度,需要将 $D_{m \max}$ 进行转换。 L_m 与其对应的地面距离 L 的关系如下

$$L = S L_m \quad (6)$$

S 为地图比例尺。Li和Openshaw在考虑原始比例尺的因素后,进行了改正^[10]

$$L = S_d L_m (1 - S_s / S_d) \quad (7)$$

其中, S_s 表示源地图的比例尺因子; S_d 表示目标地图的比例尺因子。

由式(5),式(7)可知,基于SVO用实地尺寸表示的最大网眼密度为

$$D_{m \max} = 4 / L = 4 / (S_d L_m (1 - S_s / S_d)) \quad (8)$$

部分制图专家认为图上视觉可分辨的距离 L_m 值为 $0.4 \text{ mm}^{[12]}$ 。当比例尺从1:1万综合到1:5万时,可以计算出1:5万尺度所允许的最大网眼密度

$$D_{m \max} = 4 / (50\,000 \times (0.4 \times 0.001) (1 - 1/5)) = 0.25 \text{ m}^2$$

4.2 基于制图规范确定网眼密度阈值

有些制图规范还对部分类型的网眼大小提出要求,可以转换为网眼密度阈值。例如,1:5万制图规范要求地图上街区的长、宽至少为 1.2 mm 和 1.0 mm ,道路宽度至少为 $0.5 \text{ mm}^{[13]}$ 。据此道

路中线所能包围的最小区域的长、宽分别为 1.45 mm 和 1.25 mm 。则可得出1:5万比例尺下城市道路网的 $D_{m \max}$,计算如下:

$$D_{m \max} = 2(1.45 + 1.25) \times 50 / ((1.45 \times 50) \cdot (1.25 \times 50) (1 - 1/5)) = 0.074 \text{ (m}^2 \text{ m}^2)$$

4.3 基于样图统计确定密度阈值

基于理论分析和制图规范得到密度阈值,并不适合于所有类型的网眼。对于等级中等的道路形成的网眼,其密度阈值要小于理论分析值,同时制图规范并没有相应的规定。因此,作者提出基于统计样图来确定密度阈值的方法。通过分析样图综合前后相同等级网眼密度与网眼个数的关系来确定密度阈值。

以比例尺1:1万的道路综合到1:5万为例,介绍了具体统计方法。样图分别为该数据本身(见图2(a))及已有1:5万数据(见图2(b))。其中道路等级包括主要街道和次要街道两种,则网眼分为由主要街道构成的网眼和由次要街道参与构成的网眼。计算了两种数据中网眼密度。图2(c)和图2(d)中曲线表示密度值和密度为该值的网眼数目关系,分别表示了两类网眼不同尺度下的密度分布对比。

图2(c)中的两条差异较大的曲线表示次要街道构成网眼的密度分布对比,在密度值较大的区间($0.026 \sim 0.070$),每一密度值对应的大比例尺数据中的网眼数目多于小比例尺数据中的网眼数目,表明大比例尺下很多密度值大的网眼,通过删除路段进行了合并,网眼数目变少,密度也变小;两条分布曲线在密度值较小区间($0.010 \sim 0.026$)则相反,这是大比例尺下网眼被合并后,小密度网眼增多的结果。密度值 0.026 m^2 是两个区间的分界线,可以肯定密度大于 0.026 的网眼在小尺度下需要选取。然而由于小尺度中小密度网眼的增多,密度阈值要小于该值,但相差不会太大,可以近似认为 0.026 是1:5万比例尺下次要街道网眼要求的密度阈值。

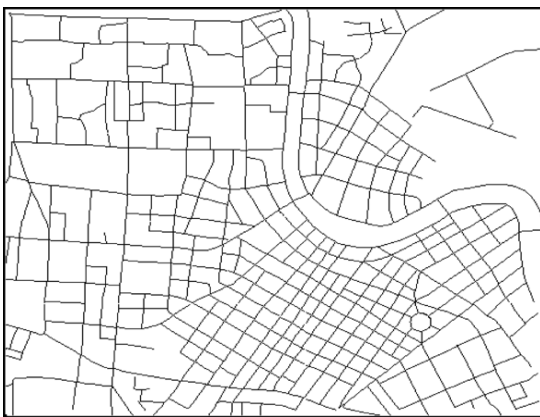
从图2(d)的分布曲线可知,两种主要街道网眼密度分布几乎吻合,表明主要街道可以直接在1:5万数据中表示,几乎没有舍弃。但已知其最大密度值不能超过 0.074 ,密度阈值设为该值。

可见,有时对等级较高的道路很少进行删除,以PLD或从规范获取值为其网眼密度阈值;对等级中等的道路如次要街道、乡村路等组成的网眼,要进行密度阈值的统计;等级更低的一些道路如

内部道路等组成的网眼,密度阈值为0,表明在小尺度数据中直接删除所有路段。所得系列密度阈值,对设置其他数据从1 1万综合到1 5万时的密度阈值,具有参考作用。密度大于阈值的网眼就是需要选取的网眼,但舍弃哪些路段,还要由其他参数和制定规则来决定。



(a) 1 1 万道路网



(b) 1 5 万道路网

(c) 次要街道网眼密度分布对比

(d) 主要街道网眼密度分布对比

图2 网眼密度阈值的估算

Fig.2 Mesh density threshold computation

5 路段取舍过程描述

5.1 路段取舍参数

文献[7]统计了道路选取对道路的等级、长度、车道数目、道路允许的交通方式数目、宽度、及连通性等属性的依赖程度,其比重分别为27%,20%,18%,15%,13%,7%。然而,本文道路的连通性通过其他途径来解决,一般的数据也很难提供交通方式和车道数目。对于其他参数,在提供不完整,或者在比较后差别微小时,得出的选取结果不能充分顾及道路在几何、语义方面的特性。为此,本文增加了两个参数,一是考虑视觉的道路长度(stroke长度);二是stroke的度。它们结合了路段之间在几何、语义及拓扑方面的特性。

人工选取道路时,道路的视觉长度对道路取舍起着重要作用。Thomson和Richardson^[6,7]提出了道路与水系等网络结构中的stroke概念,其来源于Gestalt视觉感知中的“良好连续性(Good Continuation)”定律,即“表现为相同方向的那些元素容易被看成连在一起构成一个stroke”。这里方向相同是指相接于同一结点的路段在该点处方向一致,如图3所示,6条路段构成了3个stroke。本文建立stroke时,不仅考虑了路段名称,还考虑了方向一致,且名称优先。stroke长度模仿了人工选取道路时,人眼对路段所在道路的长度判断,结合了几何与语义属性来反映路段的重要性,可以替代道路长度来使用。

图3 路段与stroke

Fig.3 Road segments and its strokes

另外一个重要参数是一个stroke包含的路段个数,类似于图论中度的概念,这里称为stroke的度,用符号 D_s 来表示。一般来说,stroke的度越大表明道路与更多的道路相互交叉,道路也就越重要。当stroke的度为1时,表明一条路段在几何视觉及名称语义上与其他路段关联最小,在选取中最容易被舍弃。

在选取中,发现所用参数的优先级明显,如在网眼边界路段中只存在一条低等级的路段,无疑就舍弃这条路段,故路段重要性判断采用了按参

数的优先级来进行。按照从重要到次要的优先级顺序,所采用的参数为:道路的等级(C)、stroke 的度(D_s)、stroke 长度(L_s)及路段的长度(L)。这些参数完全可以将每个网眼上最次要的路段筛选出来。其他一些隐含的属性,如路段与码头等重要地物相连,路段将不被舍弃,选取中也应被考虑。下面使用这些参数建立了规则,应用于路段取舍。

5.2 渐进式确定网眼边界上的舍弃路段

通过循环剥离密度最大网眼,渐进式确定网眼上的舍弃路段并与邻接网眼合并,图4和表1给出了其基本过程。

图4 网眼 M_c (见图1)及 M_c 上路段取舍与网眼合并
Fig.4 The elimination of road segments and merging of meshes into adjoining meshes

对于每个密度大于阈值的网眼,确定舍弃路段的方法,是根据参数的优先等级逐步缩小路段范围,直至找到最次要路段,并确定是否被舍弃,

称为渐进式筛选法。本文制定了相应的取舍规则,它们简单、易于实现,按顺序描述如下:

规则1:从等级最低的路段中选择要舍弃的路段;

规则2:如存在stroke 的度 D_s 为1的路段,舍弃其中长度 L 最小的路段;

规则3:如果没有 D_s 为1的路段,舍弃stroke 长度 L_s 最小的路段,当长度 L_s 差值不大时,舍弃长度 L 值最小的路段;

规则4:与码头等重要地物相连的路段不舍弃;

规则5:所在的stroke 长度达到某一长度时不舍弃。

其中,规则1,规则2,规则3为删除性规则,基于这些规则找出最次要的路段,如表1中网眼 M_c 边界上路段,通过规则1,规则2得出a被舍弃;而规则4,规则5为保留性规则,防止了在其他属性方面重要的路段被舍弃。这说明对于需要选取的网眼,并非必须舍弃路段,需要全面考虑路段的几何与语义特征,确保所选取的为重要道路,舍弃的为次要道路。

悬垂路段的取舍规则相对简单。类似数据库SQL语言,运用逻辑运算组合参数作为取舍规则。比如,规定路段的stroke 的度(D_s)为1同时stroke 长度(L_s)小于100 m的需要舍弃。当然,用户可以根据需要定制不同的规则。

表1 网眼 M_c 及 M_c 边界上路段属性及舍弃路段的确定

Tab.1 Parameters of road segments on the boundary of mesh M_c and M_c and to confirm the road segments to be eliminated

网眼ID	路段ID	路段等级	Stroke 度	Stroke 长度/ m	路段长度/ m	其他属性	删除标识	
M_c	a	次要街道	1	88.8	88.8	无	x	
	b	次要街道	3	329.5	74.3	无		
	c	次要街道	15	934.2	85.4	无		
	d	次要街道	2	156.7	70.4	无		
M_c	b	e	次要街道	3	329.5	150.5	无	
		c	次要街道	15	934.2	85.4	无	
	d	d	次要街道	2	156.7	70.4	无	x
		f	次要街道	2	156.7	85.8	无	x

5.3 循环剥离密度最大的网眼与邻接网眼合并

舍弃路段的网眼必须是从密度最大的网眼开始,以防止网眼被过度合并,造成不必要的路段删除。因此,要对密度大于阈值的网眼进行排序,剥离出密度最大的网眼,如果舍弃路段,则与邻接网

眼合并,合并后的网眼密度仍可能大于阈值,还要再次进行排序,形成循环处理的过程。网眼合并的主要步骤如下:

步骤1:对密度超过阈值的网眼按密度值从大到小排序,剥离出密度最大的网眼;

步骤2:比较道路网眼边界上各路段的重要性,判断最次要的路段,如需舍弃,删除并作标识;

步骤3:根据标识路段的左右多边形拓扑关系,合并网眼并重新组织路段,标识网眼类型,如果该网眼密度超过阈值,加入密度超过阈值的网眼集,并排序;

步骤4:依次从网眼集剥离出网眼,按2、3步处理,直至网眼集中所有网眼被处理。

在步骤3中重新组织路段是一种虚拟的过程。例如图4中合并后的网眼 M_c ,路段 d 由路段 d 和 f 组成, b 由路段 b 和 e 组成,判断舍弃路段时,仍以5条路段 b, d, c, e, f 来处理,确定出舍弃路段为 d ,要形成网眼的合并,同时也将路段

f 舍弃,即舍弃重新组织的路段 d 。这样逐步地合并网眼,保证了道路网在选取后的拓扑连通。

6 实验与结论

对上文提到的1:1万道路数据进行了综合到1:5万的选取实验。密度阈值估算见第4节,选取过程见第5节,其中为悬垂路段设置的规则为:删除图上stroke长度小于1mm且为次要街道的悬垂路段。图5(a)为选取结果;图5(b)为保留与舍弃的道路对比;图5(c),图5(d)分别为选取结果与已有图相互差异的道路。选取结果中保留与舍弃的道路以及两种原始数据中的道路长度统计见表2。

(a) 选取结果 (b) 保留与舍弃道路的对比 (c) 已有图比本文选取结果多出的道路 (d) 选取结果比已有图多出的道路

图5 选取结果及比较

Fig.5 Comparison between automated results and existing map

表2 选取结果统计与对比

Tab.2 Statistics about result of automated result and existing roads / km

	1:1万道路数据	1:5万	
		自动选取结果	已成图数据
花费时间		1 min	
总长度	147.4	120.4	120.7
舍弃的路段总长度		25.6	
舍弃的悬垂路段长度		1.4	
与自动选取的差异			4.7
与已有图的差异		5.0	
相同路段长度		115.7(96%)	115.7(96%)

从图5及统计数据,可以得出:

1. 自动选取结果与已有图数据相同率达到了96%,说明选取效果较好;

2. 自动选取结果仍然是连通的,且密度差别保持的很好;

3. 相比手工选取,效率更高。人工进行选取时,为保持连通性和密度差别,需经常查看路段属性,效率低,也容易出错。

此外,作者还对江苏省16幅1:1万的道路数据(总长度为1538.5公里)综合到1幅1:5万时进行了选取实验,选取结果与已有图相同率达到93%,效果良好。

综上所述,基于网眼密度的道路选取方法,选取结果不仅很好地保持了道路网在拓扑、几何与语义方面的重要特征以及密度差别,而且效率高。下一步的研究是将道路选取与其他要素如居民地等联系起来考虑,使道路选取更加合理。

参考文献:

- [1] T. PFER F, PILLEWIZER W. The Principles of Selection: a Means of Cartographic Generalization [J]. The Cartographic Journal, 1966, 3(1): 10-16.
- [2] MACKANESS W A, BEARD M K. Use of Graph Theory to Support Map Generalization [J]. Cartography and Geographic Information Systems, 1993, 20(4): 210-221.
- [3] MACKANESS W A. Analysis of Urban Road Networks to Support Cartographic Generalization [J]. Cartography and Geographic Information Systems, 1995, 22(4): 306-316.

- [4] JIANG B, CLARAMUNT C. A Structural Approach to the Model Generalization of an Urban Street Network[J]. *GeoInformatica*, 2004, 8(2): 157-171.
- [5] JIANG B, HARRIE L. Selection of Streets from a Network Using Self-organizing Maps[J]. *Transactions in GIS*, 2004, 8(3): 335-350.
- [6] JO O E. Causes and Consequences of Map Generalization [M]. London: Taylor & Francis, 1998.
- [7] LI Z L, CHOI Y H. Topographic Map Generalization: Association of Road Elimination with Thematic Attributes [J]. *The Cartographical Journal*, 2002, 39(2): 153-166.
- [8] THOMSON R C, RICHARDSON D E. The 'Good Continuation' Principle of Perceptual Organization Applied to the Generalization of Road Network [A]. *Proceedings of the 19th International Cartographic Conference [C]*. Ottawa: [s. n.], 1999. 1 215-1 223.
- [9] THOMSON R C. The 'Stroke' Concept in Geographic Network Generalization and Analysis [A]. *Proceedings 12th International Symposium on Spatial Data Handling [C]*. Vienna: [s. n.], 2006. 681-697.
- [10] LI Z L, OPENSHAW S. Algorithm for Automated Line Generalization Based on a Natural Principle of Objective Generalization [J]. *International Journal of Geographical Information Systems*, 1992, 6(5): 373-389.
- [11] LI Z L, OPENSHAW S. A Natural Principle for Objective Generalization of Digital Map Data [J]. *Cartography and Geographic Information System*, 1993, 20: 19-29.
- [12] MILLER J. Optimum Point Density and Compaction Rates for the Representation of Graphic Lines [A]. *Proceedings AUTO CARTO8 [C]*. Baltimore: [s. n.], 1987. 221-230.
- [13] State Bureau of Surveying and Mapping. *Compilation of Standards for Mapping and Surveying — Map Making and Printing Volume [M]*. Beijing: Standard Press of China, 2003. (国家测绘局. 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.)

(责任编辑: 雷秀丽)

勘误:

1. 本刊2007年第2期第181页, 作者简介中“毛政元(1964), 男, 湖北武冈人”, 应为: “毛政元(1964), 男, 湖南武冈人”。特此更正, 并向作者和读者致歉。

2. 本刊2007年第二期第184页中图4有误, 现更正如下图, 特此说明, 并向作者和读者致歉。

Fig.4 The stripe corresponding to the peak value of point density