利用车辆轨迹数据提取道路网络的几何特征与精度分析

蒋益娟,李　响*,李小杰,孙　靖
(华东师范大学 地理信息科学重点实验室,上海 200062)

摘要: 电子地图数据是各种智能交通系统的基础数据,对方便人们交通出行、解决城市交通拥堵问题具有重要的意义。其中,道路网络数据又是电子地图的重要组成部分。传统的道路网络数据更新方法往往需要耗费大量人力物力,因此,从其他数据源中如遥感影像数据、LIDAR 数据等提取道路网络的相关研究,已成为国内外的研究重点。车载定位设备的广泛应用使车辆轨迹数据的大量获取成为可能,轨迹数据是对车辆行驶路径的完整记录,同时也是道路网络几何特征的直接反映。当轨迹数据量足够大时,则可利用其构建路网,用于更新或修正现有地图上的路网空间信息。针对车辆轨迹数据的特点和道路网络的特性,本文提出一种细化的道路网络几何特征提取方法。车辆轨迹数据是矢量数据,将其转换为栅格数据后,就可采用图像细化的方法处理。图像细化可以在保持原图像拓扑结构不变的情况下,快速地提取出图像的中心像元,并且有效去除冗余信息。本文以上海陆家嘴的车辆轨迹数据为例进行了实验,结果表明,利用车辆轨迹数据构建路网不仅可行,而且简单、高效,取得了良好的效果。

关键词: 车辆轨迹;栅格化;细化;道路网络
DOI: 10.3724/SP.J.1047.2012.00165

1　引言

电子地图是各种智能交通系统的基础数据,对方便人们交通出行、解决城市交通拥堵问题具有重要的意义。其中,道路网络数据又是电子地图的重要组成部分。尽管目前已经存在数字化的大量道路网络空间数据,但其更新速度往往落后于现实生活中道路网络的变化,如原有道路的废弃,新建道路的开通等。而车辆轨迹数据是对车辆行驶路径的完整记录,同时也是道路网络几何特征的直接反映,因此,利用车辆轨迹数据提取道路网络的几何结构,可实现对道路网络空间数据的更新。目前,测绘或交通研究部门已经开始利用一些专门的数据采集车(Probe Vehicle)搜集路网几何特征,但由于其数量及工作时间有限,使其时空覆盖范围均受到影响,降低了其实用性。而目前大量存在的出租车轨迹数据恰恰可以弥补这一不足。

道路网络几何特征的获取途径,可以分为直接测量和间接获取两大类。直接测量即通过测量的方法获得道路网络几何特征点空间位置,包括传统的测量方法、GPS 测量方法等。直接测量的方法往往需要耗费大量的人力物力。间接获取的方法是指从其他类型的数据源获取道路网络的几何特征要素的空间位置的方法。目前,常见的间接获取方法包括从遥感影像数据中提取、从 LIDAR 数据,以及车辆轨迹数据提取等。遥感影像数据和 LIDAR 数据都属于遥感数据,而车辆轨迹数据是由车载定位设备如 GPS 获取的数据。利用轨迹数据获取路网空间信息通常是以连接轨迹点构成轨迹线,这种方法从海量车辆轨迹数据中提取路网存在一定的难度,如车辆在道路网络的某些区域有较大的重复或者交叉,这些重复或者交叉的地方很难进行取舍。为此,本文提出一种以图像细化算法的路网构建方法,其不但可保持原图像拓扑结构不变且有效地提取出

收稿日期: 2011-03-09; **修回日期**: 2012-01-14.
基金项目: 国家自然科学基金项目(J0730534);上海市自然科学基金项目(11ZR1410100);上海市城市化生态过程和生态恢复重点实验室开放基金,地理信息科学教育部重点实验室主任基金。
作者简介: 蒋益娟(1988-),女,硕士研究生,研究方向:交通地理信息系统。E-mail:jyj2166828@126.com
*通讯作者:李响,男,博士,教授,研究方向:交通地理信息系统、空间数据挖掘及空间优化配置算法等。
E-mail:xli@geo.ecnu.edu.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

道路中心线，并可有效去除大量冗余信息。

图像细化主要应用于指纹纹线、汉字或字母、医学影像等的细化。常用方法如 Hilditch 经典细化算法[1]、Rosenfeld 并行细化算法[2]、Pavlidis 异步细化算法[3]、Zhang 快速并行细化算法[4-7]、数学形态学细化算法[8]、图像边缘搜索编码细化算法[9]等。这些算法往往与特定的应用相结合，其用于轨迹数据的可行性及适用性是本文的研究目的之一。

2 车辆轨迹数据构建路网的原理及技术路线

车辆轨迹数据物理上由大量轨迹点构成，每个轨迹点均包含车辆标示信息、该点经纬度及该点采集时间等信息。利用轨迹数据获取路网空间信息通常是以连接轨迹点构成轨迹线，再从大量轨迹线中抽取出道路中心线的方法实现。当轨迹线数量较小，路网结构简单的情况下，这是一种行之有效的方法。但数据量小同时也意味着其空间覆盖范围有限，或者说其能够覆盖整个路网的时间周期较长，令其失去了可快速获取路网更新信息的优势。另一方面，当轨迹线数量较多时，虽然在空间上满足了充分覆盖的要求，但线与线之间的复杂交叉必将令自动判断路网几何特征变得异常困难。如图 1 所示，其中三条轨迹线可能在同一个路段上多次相交，产生多个疑似交叉口，令对真正道路中心线及交叉口的判断变得异常复杂。

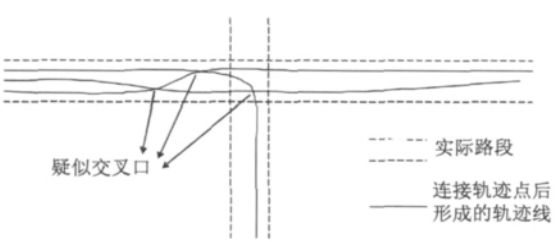

图 1 利用轨迹线方法提取路网几何特征的复杂性

Fig. 1 The complexity of extracting geometrical features of road networks by linking trajectory points

当将所有轨迹点均显示于二维平面上，并调整代表轨迹点的图斑大小后，则可从视觉上看出构成轨迹数据所在的路网空间构成。通过将上述轨迹点图在给定比例尺下转成栅格数据，则可获得构成交通网络的路段表面。进而利用图像细化算法能够保持原图像拓扑结构不变的特点，快速地提取出图像的中心像元，并有效去除冗余信息。最后，再通过矢量化方法将细化后的栅格数据结果转换成矢量数据，即，构成道路网络的线段，用于空间分析，其技术路线如图 2 所示。

首先，显示研究区内所有轨迹点，针对每个轨迹点建立缓冲区并调整缓冲区半径，确保图中属于同一路段的大部分轨迹点在空间上联成一体，缓冲区半径需要配合实际轨迹点的数量进行相应调整，同一路段上轨迹点间的距离与缓冲区半径成正比；之后，利用栅格化方法将轨迹点图转换成栅格数据，令每个栅格均具有两种取值，即“有轨迹点数据”或“无轨迹点数据”；栅格数据中，部分取值为“有轨迹点数据”的栅格实际是由于定位误差造成的，因此，需要剔除，在具体选择时，我们可以计算取值为“有轨迹点数据”的邻接栅格数量，当数量小于一个给定值时，则认为其与整个路网无连通关系，将其修正为“无轨迹点数据”；然后利用细化算法获得路网骨架信息；最后利用矢量化方法获得构成网络的线段信息构建道路网络。

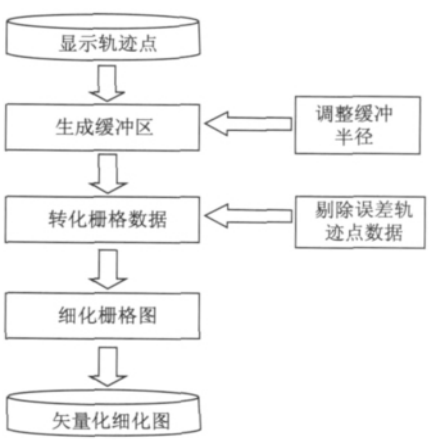

图 2 技术路线框图

Fig. 2 Overview of the proposed approach

3 图像细化算法的选择

3.1 细化算法的定义

细化算法旨在保持图像几何结构和拓扑结构的不变，以快速提取图像的中心像元并且减少图像的冗余信息，降低图像处理的复杂程度，节省内存，以提高图像处理的速度和效率。一般来说，图像细化的定义是令 R 为一平面集，B 是其边界，且 P 是 R 中的一点。在 B 上 P 的一个最近的近邻是 B 中的一点 M，使得在 B 中没有其他点与 P 之间的距

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

离小于 PM,如果 P 不止有一个最近的近邻,那么就称 P 为 R 的骨架点,称所有骨架点的并为 R 的骨架或中轴。像元集 R 的骨架是按下述步骤求出的集:首先,确定骨架像元和 R 的轮廓像元;然后,移去所有不是骨架的轮廓像元,并且用这样求出的集代替 R,重复这一过程,直至剩下仅由骨架像元所组成的集为止,称之为细化[10],如图 3。



图 3　图像细化示意图

Fig. 3　Map of thinning of the image

3.2　典型细化算法的描述

细化处理一般是针对二值图像的,细化算法应该满足如下基本要求:(1)保持原图像的拓扑结构和几何结构不变;(2)保持原图像的连通性;(3)尽量位于原图像的中心线;(4)单像元宽[11];(5)尽量保留特征点,如交叉点;(6)在保证细化效果的前提下,尽量达到快速细化[12]。

按照细化的最终结果产生的过程,细化算法大致可以分为 3 类:(1) 非迭代细化算法:直接通过距离的度量来获取骨架。(2)迭代细化算法:基于像元的邻域分析(如八邻域)或者模板匹配来获取骨架[13]。(3)数学形态学细化算法:由数学形态学中的基本运算膨胀和腐蚀组合变换求得图像的骨架像元[8]。

如上文所述,虽然已经出现了种类繁多的细化算法,但每一种算法的出现往往与特定的应用相结合,如某些细化算法适合于面状要素较多的图像,而有些算法适合于去除离散的噪音图斑。针对同一幅图像,采用不同的细化算法,图像的细化效果也各不相同。因此,我们希望从现有的具有代表性的细化算法中选择出适合这类应用的算法。非迭代算法通过计算所有边界点到内部点的距离来直接确定中心线,特别是当点的数量很多时,中心线确定难度大且计算复杂,所以,非迭代算法在图像细化处理中的应用效率不高。迭代细化算法中,串行算法比并行算法一般能够更加保证单像元宽,但是,在细化速度等方面不如并行算法。数学形态学细化算法具有算法简单,处理速度快,所以,在数字图像处理中广为应用。所以,本文拟采用典型的迭代细化算法和数学形态学算法进行道路网络细化。迭代细化算法包括 Hilditch 细化法、Zhang 快速并行细化法、Deutsch 细化法[14]、交叉数细化 1 和交叉数细化法 2(考虑对称性)[15]。其中,除 Hilditch 经典细化算法是串行算法外,其余算法皆为并行算法。

3.3　典型算法的细化应用分析

我们分别实现了 Hilditch 经典细化算法、Zhang 快速并行细化算法、Deutsch 细化算法、交叉数细化法 1、交叉数细化法 2(考虑对称性)及数学形态学细化算法。

上述 6 种细化算法被应用于同一组数据,其局部结果如图 4 所示,细化时间如表 1 所示。

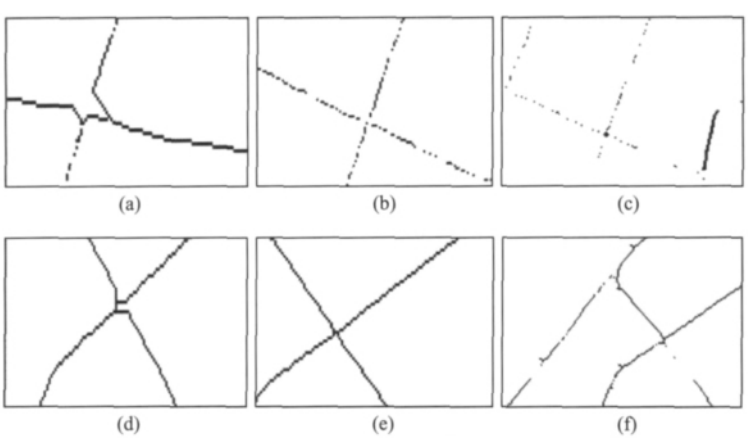

图 4　车辆轨迹细化图,其中,(a)Hilditch 法细化图,(b)Zhang 快速并行法细化图,(c)Deutsch 法细化图,(d)交叉数法 1 细化图,(e)交叉数法 2 细化图,(f)数学形态学法细化图

Fig. 4　Thinning results of vehicle trajectory points based on a: Hilditch method; b: Zhang method; c: Deutsch method; d: switching-figure method 1; e: switching-figure method 2; f: mathematical morphology method

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 1 各种算法的细化时间比较(秒)

Tab. 1 Comparison of computational time with different thinning methods (s)

细化算法	细化时间
Hilditch 细化法	38.38
Zhang 快速并行细化法	51.37
Deutsch 细化法	39.9
交叉数细化法 1	23.45
交叉数细化法 2	23.39
数学形态学细化法	36.18

通过上述几种算法的比较可以得出:(1)从细化时间来看,Zhang 快速并行细化法最慢,交叉数细化法 2 最快。(2)从细化效果来看,Hilditch 法细化图,交叉数法 1 细化图和交叉数法 2(考虑对称性)细化图总体图像清晰,符合单像元性要求,但是 Hilditch 法和交叉数法 1 细化图在交叉点位置发生了偏移,而交叉数法 2 克服了交叉口偏移的问题,达到了很好的细化效果;Zhang 快速并行细化法和 Deutsch 细化法的细化图的道路连通性差,特别是 Deutsch 细化法细化图,不仅图像不清晰,而且还存在算法不收敛和图像局部区域细化不完全等问题;数学形态学细化法细化图最大的不足就是细化后产生的毛刺,不能准确反映路网结构的几何特征。因此,我们选择交叉数细化法 2 作为应用于车辆轨迹点栅格数据的细化算法。

4 车辆轨迹数据提取路网的应用与案例分析

实验在 Windows XP,配置为 T5500 1.66GHz 0.99G 内存的环境下,以 Microsoft Visual Studio 2005 C# 为工具开发了细化算法,并利用 ArcGIS 软件实现了其他步骤,原始轨迹点数据存在于 shapefile 格式的文件中。本文的数据来源于 1000 辆出租车连续一周的轨迹点信息,其数据采集频率为每 10 秒 1 次,研究区为上海陆家嘴,涉及轨迹点数量为 3435 个,其在 ArcGIS 中的显示结果如图 5 所示。

根据车辆轨迹点的空间分布,我们将每个轨迹点建立了一个半径为 18.1 m 的缓冲区,作为代表该缓冲区的图斑,之后,利用 ArcGIS 软件,将此矢量图斑文件转化成栅格数据,如图 6 所示。该栅格数据中,每个栅格的取值情况如下:凡是被图斑所覆盖区域即取值为"有轨迹点数据"(图中以黑色表示),否则取值为"无轨迹点数据"(图中以白色表示)。

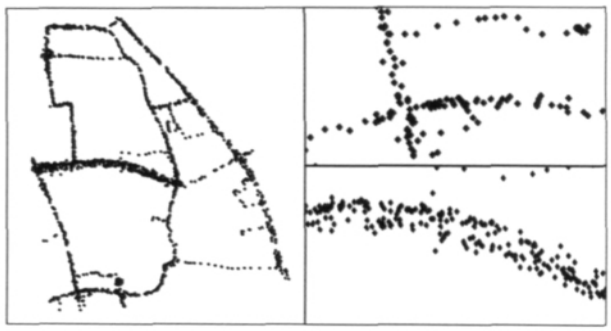

图 5 车辆轨迹点矢量图

Fig. 5 Positioning points of vehicle trajectory

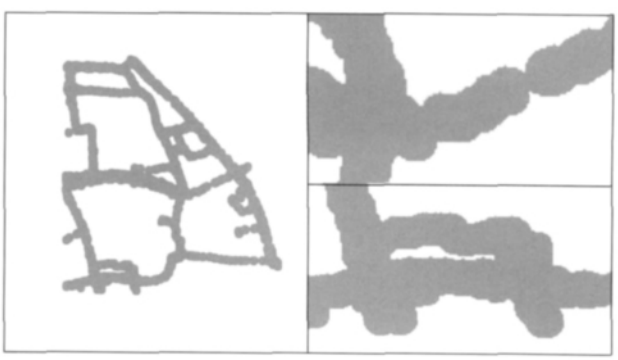

图 6 车辆轨迹点栅格图

Fig. 6 Rasterized vehicle trajectory points

之后,利用交叉数细化法 2 执行细化操作,其结果如图 7 所示。



图 7 车辆轨迹细化图

Fig. 7 Thinning results of rasterized vehicle trajectory points

此细化结果图可直接利用 ArcGIS 软件提供的矢量化工具转化成线形矢量文件,通过坐标转换和配准,即可得到矢量化的道路网络几何数据。为验证结果,我们将上述利用车辆轨迹获得的路网数据

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

与来自于上海市测绘院 2007 年发布的电子地图进行比对,如图 8 所示(其中,紫色线段表示利用车辆轨迹数据构建的路网,而蓝色表示来自于电子地图的道路网络)。

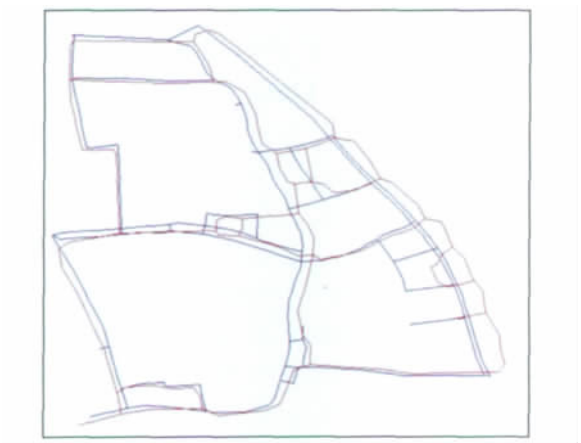

图 8　车辆轨迹图和电子地图的比较

Fig. 8　Comparison of existing road network and the network extracted with thinning methods from vehicle trajectory points

由图 8 可以看出,从精度来看,利用车辆轨迹数据提取的道路网相对于现有电子地图整体向右下方偏移,这主要是因为轨迹数据所采用的 WGS84 坐标和正式电子地图的坐标系统不一致。从整体来看,利用车辆轨迹数据提取的道路网和现有电子地图结构基本一致,但是,通过局部的比较可以发现,在研究区的中部,车辆轨迹道路网和现有电子地图的路网结构不一致,经过实地调查发现,其系因该区道路重新布局,而电子地图尚未更新所致。此外,图 8 中两种路网数据比较也不难发现,通过轨迹数据建立的路网在线段光滑程度上不及电子地图中的数据,这主要是由于矢量化过程造成的线段锯齿。

5　结语

随着车载定位设备的普及应用,车辆轨迹数据的来源越来越丰富,尤其是出租车的轨迹数据具有覆盖面广、信息更新速度快等特点。车辆轨迹数据是车辆在沿着道路行驶的过程中利用车载定位设备获取的,是道路网络几何特征的直接反映,因此,可以作为构建路网几何结构的信息来源。利用车辆轨迹构建的道路网络,可更新现有电子地图的道路信息,对电子地图的错误信息进行更正。针对车辆轨迹数据存在数据量大及误差与信息并存等问题,本文提出了一种基于细化算法的路网构建方法,说明了该方法的原理和主要流程,分析了不同算法的优缺点,并从中选择出交叉数细化法 2,其在细化车辆轨迹图时,不仅可以快速地细化车辆轨迹图,而且细化后的图像光滑,连通性好,图像清晰且符合单像元性要求,也不存在断点、毛刺、细化不彻底等问题,是细化车辆轨迹栅格图的有效方法。最后以上海陆家嘴出租车轨迹数据为例进行试验,结果表明,利用车辆轨迹数据构建路网取得了良好的效果。

在今后的研究中,我们将在细化过程中考虑高程信息,用以区别处于不同平面的路段空间信息。此外,从轨迹数据中还可以提取路网的拓扑特征,这一点也将是我们有待深化研究的重点。

参考文献:

[1] 刘利娜, 吴锡生. 提高手写体数字细化效果的改进算法[J]. 计算机工程与应用, 2010, 46(3): 137-139.

[2] Stefanelli R, Rosenfeld A. Some parallel thinning algorithms for digital pictures[J]. Journal of the ACM, 1971, 18(2): 255-264.

[3] Pavlidis T, Ali F. Computer recognition of handwritten numerals by ploygonal approximations[J]. IEEE Transactions, System, Man and Cybernetics, 1975, 5(6): 610-614.

[4] Zhang T Y, Suen C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27: 236-239.

[5] 吴选忠. Zhang 快速并行细化算法的扩展[J]. 福建工程学院学报, 2006, 4(1): 89-92.

[6] 杨容浩, 余代俊. 扫描地图自动矢量化若干细节问题的解决方法[J]. 四川测绘, 2008, 31(3): 127-130.

[7] 张学东, 张仁秋, 关云虎. 一种快速的手写体汉字细化算法[J]. 计算机应用与软件, 2009, 26(11): 17-19.

[8] 张英琦, 张庆林. 数学形态学应用于二值图像的细化[J]. 焦作工学院学报, 1997, 16(4): 8-43.

[9] 尹铁源. 机械手性能标定直线运动轨迹的特征检测技术[D]. 沈阳工业大学, 2008, 1-79.

[10] 帕夫利迪斯(著), 吴成柯(译). 计算机图形显示和图像处理的算法[M]. 北京: 科学出版社, 1985, 1-403.

[11] 施启乐. 反求系统中图像获取与处理若干问题的研究[D]. 华中科技大学, 2004, 1-55.

[12] 崔凤奎, 王晓强, 张丰收. 二值图像细化算法的比较与改进[J]. 洛阳工学院学报, 1997, 18(4): 48-52.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

[13] 张昊，徐刚. 基于四邻域的二值图像细化算法[J]. 信息技术与信息化，2004，(6)：24-27.
[14]刘明艳. 边缘的检测与细化研究[D]. 曲阜师范大学，2007，1-55.
[15] 李兰友. Visual C＃图像处理程序设计实例[M]. 北京：国防工业出版社，2003，1-305.

Geometrical Characteristics Extraction and Accuracy Analysis of Road Network Based on Vehicle Trajectory Data

JIANG Yijuan, LI Xiang, LI Xiaojie and SUN Jing
(*Key Laboratory of Geographical Information Science, East China Normal University, Shanghai* 200062, *China*)

Abstract: Electronic map data is of enormous importance for analyses of travel, urban congestion, etc. It is growing an important basic data of transportation research and widely employed in vehicle navigation system and intelligent transportation system. Of electronic map data, road network data is a fundamental part and often updated. As a result of a lot of expense of manpower and resources of traditional method, more and more research efforts have been made on extracting geometrical characteristics of road network from other data resources like LIDAR data. The wide use of mobile positioning devices makes it possible to obtain a large volume of vehicle trajectory data within a short period. These data are complete records of vehicles' movement and able to reflect geometrical characteristics of road network as well. By this means, vehicle trajectory data can be employed to figure out changes happening to the spatial extent of road networks as long as enough data are available, and then we can update the road network database. Since trajectory data integrate both spatial and temporal information and their volume is often huge, it is not straightforward to realize the aforementioned target. In view of this, the paper proposes a thinning—algorithm—based method to construct road network. Vehicle trajectory data are vector data, are first converted to raster data, then thinning method can be used to extract geometrical characteristics of road network. With this method, center lines of road networks can be extracted, network topology be maintained, and meanwhile redundant information be removed. A case study in Lujiazui, Shanghai with taxi trajectory data is conducted and the results are evaluated.

Key words: vehicle trajectory; rasterization; thinning algorithm; road network

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net