

WebGIS 矢量空间数据压缩方法探讨

李青元^{1,2} 刘晓东² 曹代勇¹

¹(中国矿业大学煤炭资源教育部重点实验室, 北京 100083)

²(中国测绘科学研究院 政府地理信息系统研究中心, 北京 100039)

摘要 对 WebGIS 在网上传输的矢量空间数据进行压缩的方法进行了研究,并在文献[1]提出的矢量空间数据压缩方法的基础上,在以下 4 个方面有了新的发展:(1)对地图坐标数据不仅可以用 int 型,甚至可以用 short 型代替传统的 double 型或 float 型;(2)得到数据映射范围的计算公式;(3)可用映射范围为虚拟屏幕,对矢量地图上弧段的型值点进行“滤点压缩”,如果“锯齿”的台阶宽度 $stepWidth=1$,图形的精度是无损的;如果 $stepWidth>1$,精度有损;(4)用“长边加点”来处理“偏移量压缩”中,弧段两相邻型值点间 x 或 y 的偏移量大于 127 的个别“长边”,在进行了一系列的压缩措施后,可以使矢量地图的数据量压缩率达 80% 以上。

关键词 WebGIS 矢量空间数据 数据压缩

中图法分类号: TN919.8 **文献标识码:** A **文章编号:** 1006-8961(2001)12-1225-05

Research in WebGIS Vector Spatial Data Compression Methods

LI Qing-yuan^{1,2}, LIU Xiao-dong², CAO Dai-yong¹

¹(The Key Laboratory of Coal Resource, Ministry of Education, CUMTB, Institute Road Ding 11, Beijing 100083)

²(Government GIS Research Center, CASM, Beijing 100039)

Abstract This paper introduces author's work in compress to vector spatial data, which will be transported in Internet/Intranet. On the basis of vector spatial data compression method, which is proposed by [1], the authors proposes next four innovation: (1) For vector spatial data, not only integer but also short data can be used to replace traditional double or float data. (2) The map data are mapped to a "virtual screen", the formula of "virtual screen" width is: $width = DispWidthPixel \times MaxEnlargeRatio / stepWidth$. (3) In the mapping process, execute "filter points compression" to pattern points of arc (including polygon's boundary arc and line's arc). In the formula [1], if $stepWidth=1$, "filter compression" has no loss in precise of map displayed, else $stepWidth>1$, it has precise loss in map displayed. (4) "Long edge adding point" to deal with exception of "Offset compress" to pattern points in arc, which offset in adjacent two points may larger than 127. After these compressions, a map vector data may be compressed to 20% or less.

Keywords WebGIS, Vector spatial data, Data compress

0 引言

矢量空间数据由于其数据量小,表达精度高,因此在 GIS(地理信息系统)软件实现中始终占有重要的地位。随着互联网的普及与应用,WebGIS 将交互式的电子地图带到了互联网上。由于互联网带宽的

限制和人们对海量地图数据快速传输、显示要求之间的矛盾,矢量空间数据的压缩方法开始受到 WebGIS 研发者的重视。

在计算机图形学界,很早就有人采用 int 或 short 型数据作为图形的坐标,而在图象压缩方面有人采用了偏移量方法。由于地理定位的精度要求比较高,经常需精确到小数点后面 7~8 位,因而在以

处理地理坐标为主的GIS领域中,一般还是将地理坐标以float或double型数据存储和运算,只是在最后向屏幕或绘图仪输出时才强制转换为int或short型.针对WebGIS应用的特点,文献[1]对WebGIS中矢量空间数据(即矢量地图数据)压缩方法提出了以下的思路:(1)将double型或float型的坐标映射(不是强制转换)为int型坐标;(2)对每一条弧段(包括多边形边界和线状地物)只用两个int型数据记录其起点坐标 (x, y) ,其后续点坐标用相邻两点间 x, y 的偏移量代替,这样就可以将数据量压缩到原来的1/8左右.这些思想为WebGIS中的矢量空间数据压缩提供了很好的思路,笔者在开发“地学之窗”Web版的WebGIS软件中,采用了这些思想,并对一些问题有了新的思考与认识,在此与同行分享.

1 用short型代替float或double型

考虑到在WebGIS的数据存储与传输中,矢量空间数据的坐标不仅可以用int型,甚至可以用short型代替传统的double型或float型.将地图坐标映射为short型,其数据的表示范围是 -32768 至 -32767 ,或 0 至 65536 ,这样会不会导致地图表达精度的损失呢?以下是对这个问题的分析.

(1)用户在计算机上浏览地图,总是先显示全图,然后再逐级放大.若屏幕大小按 800×600 计算,则short表示的数据可以使 800×600 大小的全图放大80倍,而仍然没有精度损失.

(2)从另外一个角度考虑,很多WebGIS的地图空间数据是由传统的纸图通过数字化采集而来的.如果图纸的长度为1m,按1mm采集10个图面单位(数值化的精度也只能如此),那么一个图幅的图面横坐标或纵坐标为10000个图面单位.即使图幅长、宽达到6m(很少有这么大的纸图),数值化后的图面坐标为60000个单位,也小于short的标度范围65536.显然,从单个图幅显示的角度考虑,short型数据的坐标可以满足大部分WebGIS的需要.对于用经纬度表达的地图坐标,通过下面的映射变换也不会有显示精度上的损失.当一屏同时显示多个中、大比例尺的图幅(例如比例尺为1:25万,采用高斯投影),且图幅位于不同投影带时,地图投影存在“投影带裂隙”问题,这时可采用“几何变换”或“实时投影变换”解决图幅接边处的坐标连续问题^[2],然

后再向虚拟屏幕映射.

2 映射范围的确定

映射是将原来在大地坐标(或图纸坐标)系下表示的空间点坐标(用float型或double型表示)转换到一个虚拟屏幕的坐标系下.其算法与在屏幕上绘图的算法很类似,只不过不是往屏幕上绘制,而是往一个虚拟屏幕上“绘制”(实际是存储).如果映射后虚拟屏幕坐标系下空间点的坐标用short型数据表示,则虚拟屏幕最大的范围是 $-32768 \sim +32767$ 或 $0 \sim 65536$.虚拟屏幕是否需要设置到short型数据所能表达的最大范围?回答是否定的.从下一节可以看到,对地图的弧段(包括多边形边界和线状地物)进行“滤点压缩”时,滤掉点的比率与映射到的虚拟屏幕范围大小有关.从原则上说,虚拟屏幕范围越小,被滤掉的点就越多,“滤点压缩”的压缩率就越大,数据精度损失也就越大,反之亦然.当虚拟屏幕范围足够大时,“滤点压缩”不会滤掉原来的任何一个点,“滤点压缩”的压缩率为0,没有任何精度损失.因此,如何确定映射的虚拟屏幕范围是值得仔细推敲的问题.最简单的方法是以客户端屏幕上地图的显示范围(DispWidthPixel)作为虚拟屏幕范围进行映射,这样数据的压缩比非常大,当图形不放大时看不出精度损失,但放大时就会出现“锯齿”;另一个极端是将虚拟屏幕范围定位在 $-32768 \sim +32767$,若客户端屏幕上地图的显示范围(DispWidthPixel)为650,则在客户端放大100倍的情况下都还不会有精度损失(即不会出现“锯齿”).通常,系统常会加一些限制,不允许无限放大,即放大到一定倍率后就不能再放大,或调用下一级比例尺的图幅.如果最大放大倍率(MaxEnlargeRatio)远小于100(例如16),将虚拟屏幕范围定在 $-32768 \sim +32767$,会有很大程度的精度“浪费”.在带宽有限,数据传输等待时间较长的情况下,人们需要在“速度”与“精度”之间进行权衡,有时需要牺牲一定的精度以换取更快的速度,即可以允许在最大放大倍率下地图有一定的“锯齿”.因为图形的显示是以象素为基本单元,因此无论是曲线还是倾斜直线,放大到象素级别时都会显现出“锯齿”,当“锯齿”的台阶宽度(stepWidth)为1个或2个象素时,不易为人眼所察觉,当“锯齿”的台阶宽度为3个象素时,人眼就有察觉.当达到3个象素时,人眼就能看出明显的“锯齿”状台阶.如图1所

示,是全国 1:400 万图的一个局部在 16 倍的最大放大倍率下(大于 16 时调用 1:100 万图幅的数据),stepWidth 分别等于 1、2、3、4 时的弧段光滑程度。当 stepWidth 等于 1 时,在最大放大倍率下也几乎没有锯齿出现;当 stepWidth 等于 2 时也看不出明显的锯齿;当 stepWidth 等于 3 时可看出锯齿;在 stepWidth 等于 4 的情况下,曲线的锯齿已非常明显,但考虑到放大到最大倍率后即刻就会调下一级比例尺(1:100 万)的图幅,若人们能够容忍最大放大倍率下,弧段出现台阶为 4 个像素的“锯齿”(当放大倍率小于最大放大倍率时的“锯齿”台阶宽度小于 4 个像素),那么通过下面的“滤点压缩”,图幅弧段的型值点数目可能被压缩 47.5%,而由此带来图幅装载速度将近一半的提高。因此这种锯齿应该说是“物有所值”的,若人们能够容忍的最大放大倍率下的“锯齿”台阶宽度为 stepWidth,则映射的目的范围为

$$\text{width} = \text{dispWidthPixel} \times \frac{\text{maxEnlargeRatio}}{\text{stepWidth}} \quad (1)$$

式中, width 为映射目的范围的宽度; dispWidthPixel 为客户端屏幕地图显示宽度; maxEnlargeRatio 为客户端最大放大倍率; stepWidth 为最大放大倍率下锯齿台阶宽度。

映射范围高度的计算公式与式(1)相同,只需将 dispWidthPixel 换为 dispHeightPixel(客户端屏幕地图显示高度)就行。

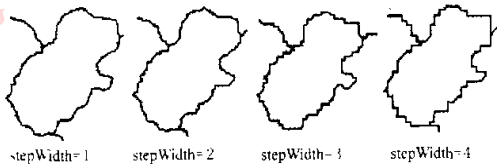


图 1 不同 stepWidth 下的曲线光滑情况

3 以映射范围为虚拟屏幕进行“滤点压缩”

映射范围确定后,可以以映射范围为虚拟屏幕进行“滤点压缩”,进行“滤点压缩”的理论依据是:原始地图上,弧段的型值点采样很密集,在其映射到所指定的虚拟屏幕后,这些型值点就更加稠密,以至于弧段上前后相随的几个点在虚拟屏幕上落在同一个像素点上;同一弧段上,这些前后相随但落在同一个

像素点上的点,只要保留一个就够了,而其余的点在极限放大范围之内,不能对图形的精度提高有任何贡献,因此应该将它们去掉,“滤点压缩”就是去掉这些多余的点,“滤点压缩”的方法很简单,只需要在映射变换过程中,在虚拟屏幕上比较要加入到弧段的型值点坐标数组中的新点是否与前一个点相同,若相同,则将新点扔掉,不相同的点才加入到弧段型值点坐标数组中,“滤点压缩”是对矢量空间数据进行压缩的有效途径之一。在式(1)中,如果 stepWidth = 1,这种“滤点压缩”是精度无损的,当 stepWidth > 1 时,“滤点压缩”则是精度有损的,人们可以根据实际应用的需要,在“精度”与“速度”之间作出权衡。

例如某全国 1:400 万专题图的所有弧段(包括多边形边界弧段和线状地物的弧段)的总点数是 598 483 个,设客户端屏幕地图显示范围 dispWidthPixel = 600; 客户端最大放大倍率 maxEnlargeRatio = 16。在最大放大倍率下弧段的锯齿宽度 stepWidth 分别等于 1、2、3、4 时,根据式(1)算出虚拟屏幕的映射范围分别是 9 600、4 800、3 200、2 400,“滤点压缩”后还剩下的点数分别是 954 561、485 234、382 087、313 924,分别滤掉了 3 922、113 249、216 396、284 559 个点,通过滤点压缩的压缩率分别是 6.7%、18.9%、36.2%、47.5%,如表 1 所示。滤点后的图形精度如图 1 所示。

表 1 不同 stepWidth 下的 width、滤掉点数及压缩率

stepWidth	width	剩下点数	滤掉点数	压缩率(%)
0.4	24 000	595 977	2 506	0.4
0.8	12 000	593 335	3 148	0.5
1	9 600	594 561	3 922	6.7
2	4 800	485 234	113 249	18.9
3	3 200	382 087	216 396	36.2
4	2 400	313 924	284 559	47.5
8	1 200	186 424	112 059	68.9
16	600	106 651	491 832	82.2
32	300	60 891	537 592	98.6

“滤点压缩”在实际往屏幕绘图时也是很有用的,在组织绘图的多边形边界或线的型值点数组时进行“滤点比较”虽然要花费一点时间,但它能减少大量的屏幕绘图操作,由于屏幕绘图操作的速度远慢于比较判断的速度,因而滤点压缩可加快图形的显示速度。

4 “长边加点”解决“偏移量压缩”中的例外

文献[1]提出的“偏移量压缩”是基于在地图的弧段(包括多边形边界或线状地物)上相邻两点坐标的 x 、 y 值在映射到虚拟屏幕后不会相差很大这样一个事实。“偏移量压缩”的思想是:地图映射到虚拟屏幕后,弧段只记录其起点(x 、 y)的坐标,而其后续型值点的坐标可用相邻两点间 x 、 y 的偏移量来表示,即弧段除起始点以外的其他后续点,用 2 个 byte 就可以表达一个坐标点。“偏移量法”将矢量地图的数据量在由 double 型映射到 int 型,数据量减少一半的基础上又减少了将近 3/4,(若由 double 型映射到 short 型,则在数据量减少 3/4 的基础上又减少了将近 1/2)。因为矢量地图最大的数据量来自于弧段(包括多边形边界和线状地物)型值点的坐标数据,因此“偏移量压缩”是矢量空间数据压缩的最有效方法,但“偏移量压缩”使用中要注意个别例外。有个别情况(例如以图幅边界为边的多边形),两点间的偏移量可能超过 127pixel。这样的边虽然只是极个别的,但如不采取特殊处理,可能会使显示出来的图形面目全非。这里提出采用“长边插点”的方法处理这些偏移量超过 127pixel 的特殊直线段。“长边加点”就是在长边上加几个点,使长边变为几个短边,每一短边的 x 偏移量、 y 偏移量不超过 127pixel。该方法很好地解决了“长边”问题。“长边加点”方法虽然增加了数据量,但需要加点的长边微乎其微,因而在映射到虚拟屏幕后,在“偏移量压缩”的同时进行“长边加点”,压缩掉的数据量远大于增加的数据量。

5 压缩效果

在以下试验例子中,原始坐标数据是 float 型,一个图幅由 5 个文件组成,包括弧段点、弧段索引、拓扑关系、线状要素、面状要素和注记等内容。总的文件大小是 1 331K。这里对其进行重组、压缩然后存为一个新文件。压缩中采取了 short 代替 float、“滤点压缩”、“偏移量压缩”等措施。压缩时采用的虚拟屏幕宽度(width)、压缩后生成的文件大小(File Size)、数据压缩率(Compress Ratio)、对虚拟屏幕宽度的可能理解($\text{Width} \times \text{Ratio} / \text{stepWidth}$)如表 2 所示。

表 2 不同 width 下,压缩后文件大小与压缩率

Width (pixel)	File Size (K)	Compress Ratio (%)	Width \times Ratio: stepWidth
300	34.0	95.95	600 \times 1.2
600	80.1	93.96	500 \times 2.2
1 200	116.9	91.22	600 \times 4.2
2 400	181.1	86.40	600 \times 8.2
3 200	217.2	83.29	600 \times 16.3
4 800	269.7	79.71	600 \times 16.2
9 600	277.8	79.14	600 \times 36.2
12 000	277.8	79.14	600 \times 40.2
24 000	278.1	79.13	600 \times 80.2

可看出,当虚拟屏幕宽度从 300pixel 增加到 24 000pixel 时,压缩后的文件大小从 54K 增加到 278K。即虚拟屏幕宽度越小,压缩后的文件越小,压缩比越大。当虚拟屏幕宽度增加到 4 800 后,压缩后的文件大小增加微小。虚拟屏幕宽度 4 800pixel,意味着当客户端屏幕显示地图宽度为 600pixel,放大 16 倍时,曲线会出现宽度为 2 个像素的锯齿台阶,如下页图 2 所示。

6 结 语

WebGIS 专家们都已认识到,在当前的网速下,WebGIS 的系统响应速度太慢,这是制约其应用的重要因素。要提高响应速度,一个重要的途径就是减少网上数据传输量,即进行矢量空间数据压缩。其有效的途径是将传统的 double 型或 float 型数据映射到 short 型的 $\text{width} = \text{dispWidthPixel} \times \text{maxEnlargeRatio} / \text{stepWidth}$ 宽度(或高度)的虚拟屏幕上,在映射的过程中对弧段的型值点进行“滤点压缩”,然后进行“偏移量压缩”,同时用“长边加点”处理个别的长边。经过这一系列的压缩后,可将图幅的数据压缩到原来的 20% 甚至更少。在数据传输到客户端后,先将“偏移量”恢复到原来的弧段或多边形,这样就不影响客户端的显示、查询等操作的效率。还有几点需要提请读者注意的是:

(1)“滤点压缩”中计算 width 时,如果 $\text{stepWidth} \leq 1$,则精度无损;如果 $\text{stepWidth} > 1$,则精度有损。

(2)“偏移量压缩”是精度无损的,“长边加点”只是为了处理个别的“长边”例外,但当虚拟屏幕范围设得过大,或图中的直线长边很多时(例如规整的公共汽车线路图、建筑平面图)，“偏移量压缩”就起不到什么压缩效果,因为过多的“长边加点”会抵消“偏

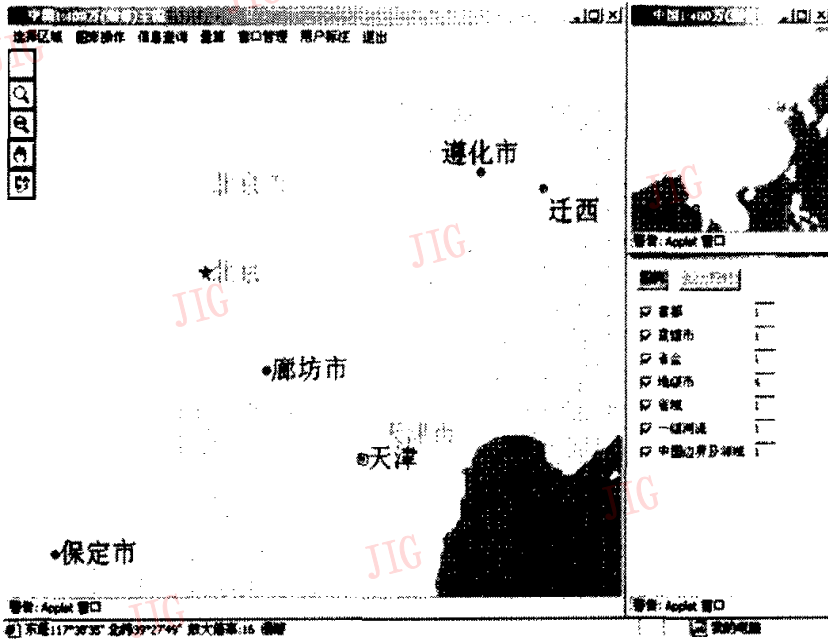


图 2 虚拟屏幕 Width=4 800pixel,客户端主图宽度 dispWidthPixel=600pixel,放大倍率 16,曲线上出现宽度为 2 个像素的锯齿台阶

移量压缩”减掉的数据量。

(3) 无论是“滤点压缩”还是“偏移量压缩”对注记、拓扑关系等数据都不起作用。

(4) 将地图坐标由 double 或 float 型映射到 short 型,对注记数据的坐标点有压缩作用,而对注记内容没有压缩作用。

参 考 文 献

- 1 李琦,杨超伟,陈爱军. WebGIS 中的地理关系数据模型研究[J]. 中国图象图形学报, 2000, 5A(2), 119~123.
- 2 常燕卿. 大型 GIS 空间数据组织方法初探[J]. 遥感信息, 2000, (2): 28~31.



李青元 1958 年生,副研究员,1982 年毕业于中国矿业大学地质系,1996 年获中国矿业大学北京校区地质系博士学位,毕业后在中国测绘科学研究院从事 WebGIS、三维 GIS 研究,现在中国矿业大学“煤炭资源教育部重点实验室”作访问学者研究。



刘晓东 1974 年生,1995 年获北京印刷学院计算机系学士学位,毕业后在辽宁日报从事信息管理工作,现为 中国测绘科学研究院硕士研究生,研究方向为 WebGIS、数据交换。



曹代勇 1955 年生,教授,博士生导师,1982 年毕业于中国矿业大学地质系,1988 年获中国矿业大学北京校区博士学位,现主要从事地学应用、构造地质学与地学计算机应用等方面的教学和科研工作。