

基于编码压缩的数字矢量地图水印技术研究^①

夏松竹^② 张健沛 杨 静

(哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

摘 要 提出了一种新的数字矢量地图信息隐藏方案。该方案通过建立动态编码压缩表缩减水印编码长度,达到提高水印容量的目的;同时将数字水印压缩码嵌入到地图属性文件的对象定义块内的冗余空间中,由于压缩信息都存储在结束符之后,使得任何地图编辑及浏览软件都无法阅读或显示嵌入冗余空间的水印信息。实验结果表明该隐藏方案鲁棒性强,具有突出的不可感知性,方案对地理精度零损伤,可抵抗剪切、旋转、投影等多种几何攻击。

关键词 数字矢量地图, 信息隐藏, 编码压缩, 数字水印, 无损

0 引 言

地图是根据属性信息定义,通过拓扑关系形成的一种图形化数字产品。数字矢量地图被广泛应用于地理信息系统、测绘科学、军事国防等多个领域^[1]。为了保护数字地图版权以及防止数据被非法窃取和滥用,研究人员提出了面向数字矢量地图的信息隐藏方案^[2]。由于攻击行为的多样性及复杂性,信息隐藏算法在实际应用中所表现的抗攻击性能同理论值存在一定偏差,水印相似度仅能达到 0.4 ~ 0.6^[3]。

目前数字矢量地图信息隐藏技术还未投入实际应用,主要原因在于现有技术无法满足数字地图的无损性要求。数字地图的无损性要求既包括图形图像等空间拓扑关系无损,也包括地理信息、地理精度数值无损。近年来,已提出很多基于空域和基于变换域的理论算法。其中,基于空域的方法^[4-6]多采用替换策略,即用隐藏信息来替换或改变载体数据^[4-6],技术上实现简单快捷,对地图的变换程度过高,使得地图的改变程度较大,严重影响数字地图的精度和图形图像特征,不具有实际应用的可行性。而基于变换域的方法^[7-10]计算复杂程度较高,对地图的精度及拓扑结构的改变幅度相对较小,但仍无法达到地图资料“零损失”的实用目标。为了解决

这一难题,哈尔滨工业大学牛夏牧教授的研究小组先后提出了面向数字矢量地图的可逆水印^[11]技术及无损信息隐藏方案^[12]。该类算法通过修改相邻坐标间的差值来嵌入水印,尽管水印的嵌入过程对地图数据仍有改变,但算法能够保证在提取水印数据时对地图数据进行真正的无损恢复,实验表明该类方法对于含有线、面图层的数字地图均取得了很好的效果。由于该方法依赖于结点坐标的相关性,所以无法适用于坐标相关性较弱的离散点图层。本文在前人工作的基础上,提出了一种新的无损信息隐藏方案,它能够大容量、无损地^[13]向数字矢量地图中嵌入水印信息。用该方案时不需要进行无损恢复,地图数据在水印信息嵌入前后无丝毫改变。另外信息隐藏普遍适用于点、线、面各基本图层,对于地图的空间拓扑结构及属性数据无明显依赖。该方案的主要技术包括:根据动态编码表及地图空间规模生成水印压缩编码;将水印编码嵌入地图内选定数字矢量对象的信息定义空间;根据地图的比例尺及基本描述信息,生成对应的密钥。

1 水印编码压缩规则

若数字矢量地图 V 的空间规模(像素总数)为 N , 二值水印 W 的容量为 $m \times n$, 则恒有 $m \times n \leq N$ 。

为了提高算法的鲁棒性,需要将 W 重复嵌入 k

① 国家自然科学基金(60873037)和中央高校基础科研业务经费项目(HEUCF100602)资助。

② 男,1973 年生,博士,副研究员;研究方向:信息安全,数据库;联系人, E-mail: xiasongzhu@hrbeu.edu.cn (收稿日期:2010-07-02)

次^[14], $k \leq \lfloor \frac{N}{m \times n} \rfloor$ 。该方法的不足是当 N 很小时, 水印仅能嵌入一次, 导致 W 由于编码过短而被轻易去除。要求 W 的编码长度不易过大。

另一方面, 过短的编码重复率较高, 同时蕴含信息量较少, 无法达到信息隐藏的应用目标。为了更好地进行版权证明, 需要提高水印编码长度来携带更多的信息, $m \times n$ 应尽可能大。

通过水印编码压缩技术^[15]能够很好地解决上述矛盾: 在合理范围 ($m \times n \leq N$) 内, 令 W 达到足够的编码长度, 通过对 W 进行编码压缩, 获得压缩编码 CW , 并向数字地图内重复嵌入 k 次, $k \leq \lfloor \frac{N}{\text{len}(CW)} \rfloor$ 。该方法既保证了水印容量, 又提高了水印的鲁棒性。

在图 1 所示的编码压缩表中, 符号“Y”为压缩码, “N”为常规编码。考虑到码元应尽量使用

水印编码压缩过程记作							
一级编码 (Y: 压缩码, N: 常规码)							
Y	N	Y	N	Y	N	Y	N
A	000000	a	011010	,	110100	6	111010
B	000001	b	011011	.	110101	7	111011
C	000010	c	011100	?	110110	8	111100
.....		:	110111	9	111101		
F	011000	f	110010	;	111000	#	111110
Z	011001	z	110011	_	111001	*	111111
扩展编码							
+	0	2	00	3	000	4	0000
5	00000						

图 1 定长编码压缩表

单一字符表示, 一级编码被设计为 6 位 2 元定长码集, 码字数为 64, 记作 $A_6 = \{A, \dots, Z, a, \dots, z, \dots, * \}$, 显而易见, A_6 的码率(信息率) $R(C) = \frac{\log_2 M}{n}$ $= \frac{\log_2 64}{6} = 1$ 。扩展编码记作 $B = \{+, 2, 3, 4, 5\}$, 扩展编码作用主要包括:

(1) 构建组合码 (BA_6), 可将水印编码动态压缩为 $k(k > 6)$ 位定长码集 A_k 。例如: $A_8 = \{(++A_6), (+1A_6), (1+A_6), (11A_6)\}$, $A_7 = \{+A_6, 1A_6\}$ 。

(2). '0' $\in B$, '0' 码作为补码, 可解决水印余码的压缩问题, 如水印 $W = \{w_1, \dots, w_i\}$, 已知 6 不被 i 整除, 根据 A_6 定义, W 可表示为 $\{w_1, \dots, w_6, \dots,$

$w_s, \dots, w_{s+s}, w_{i-j}, \dots, w_i\}$, $0 < j < 5$ 。由定义可知, $\{w_1, \dots, w_{s+s}\}$ 均可由 A_6 中码字唯一表示。而余码 $C = \{w_{i-j}, \dots, w_i\}$ 的长度不足 6 位, 无法由 A_6 中码字唯一表示。此时可在序列后 C 后扩展 '0' 码字集合 $B_{s-j} = \{0, \dots, 0\}$ ($B_{s-j} \in B$), 记作 $CB_{s-j} = \{w_{i-j}, \dots, w_i, 0, \dots, 0\}$, $CB_{s-j} \subseteq A_6$, $C = A_6 B_m$, m 为 '0' 码个数。

水印编码压缩过程记作 $CW = \text{compression}(A_6, W)$, 读作水印 W 在定长编码集合 A_6 的映射下压缩为 CW , 编码的解压缩过程记为 $W = \text{decompression}(A_6, CW, 6)$ 。

水印编码压缩技术不仅能解决水印容量与鲁棒性之间的矛盾, 而且还有两点优势: (1) 编码压缩提高了信息隐藏方案的复杂程度, 在无压缩编码集的条件下, 由数字地图直接提取到的水印编码不具有任何实际意义; (2) 提高了编码压缩技术。

2 无损信息隐藏方案

2.1 地图属性文件分析

数字矢量地图类型主要包括 MapGis、MapInfo、ArcGis、Erdas, 这几种类型的格式之间可以进行互相转换。本文以 MapInfo 的 MapX 插件^[16,17]为例, 对基于属性文件的信息隐藏方案进行分析。在 MapX 中, 主要包括 TAB 档、DAT 档、ID 档、MAP 档 4 类, MAP 文件以数据块为单位存储数据, 它主要定义了地图对象的索引、名称、坐标、资源以及删除标识等信息, 每个数据块大小为 512 字节或 1024 字节, 最新版本对每个数据块进行了 200 字节的扩展。

在二进制 MAP 文件中, 对象定义信息存储在对象定义块 (Object Definition Block, 块标记 2) 内, 包含对象类型和对象名称两部分内容, 以结束符为信息描述的结束标志, 其它字节均被空置。经实测, 地图工具软件不读取和显示结束符后的存储信息, 这为本文提出的方案提供了可行性: MAP 文件为二进制格式, 信息不易读取, 每个对象定义块都空闲了大量空间可存储水印; 当对象被删除时, 删除标记置为真, 但与之相关的属性定义等信息仍被保留。

2.2 嵌入算法

输入: 数字矢量地图 V_0 , 二值水印图像 $W(m \times n)$; 输出: 含水印的数字矢量地图 V_w 。

```

(1)  $T = \left\lfloor \frac{m \times n}{6} \right\rfloor$ ,  $S =$  地图  $V_0$  所含对象个数
(2) if  $T > S$  then 提示“水印过大,无法嵌入地图!”
(3) for  $i = 0$  to  $T$  do
(4)   计算  $W' = compression(A_0, W_i, 6)$ 
(5) end for
(6)  $L = V_0$  中属性定义块  $B[]$  的字节总长度
(7) for  $j = 0$  to  $W'[]$  do
(8)   计算  $B[j]$  有效信息的位组长度  $l$ ,
       $W'[j]$  的位组长度  $a$ 
(9)   if  $a=0$  then exit
(10)  else if  $L-1-a > l$  then
(11)  指标指向  $B[j]$  的最后一个字节  $B[j][L-1]$ 
      处,将  $W'[j]$  与  $j$  的连接串插入到数据块
       $B[j][L-1-a...L-1]$  位置处
(12) else  $j++$ 
(13) end for

```

2.3 提取算法

输入:含水印的数字矢量地图 V' ; 输出:水印 W_i 。

```

(1)  $B[] =$  按序读取文件 MAP 中块标记为 2 的所有
    对象定义块,  $S = V'$  所含对象个数
(2) 本文选用的是 MapX500 版本,其定义块大小为
    1024 字节,因此,将块的最后一个字节转化为整
    型  $int(B[][1023])$ 
(3) for  $i = 0$  to  $S - 1$  do
(4)   if  $int(B[][1023]) \neq 0$  then
(5)     { 将  $B[j][1023]$  按递增插入序列  $N[k_i]$ ,
         $k_i$  是排序后插入值所处位置
(6)       for  $j > 0$  do
(7)         if  $B[j][1023-j]$  it not null & 不是定
            义结束符 then  $j++$ 
(8)         else  $B[j][1022...1023-j]$  插入序列
             $V_w[k_i]$ , 根据  $k_i$  调整序列  $V_w[]$ 
(9)       end for
(10)    else  $i++$ 
(11)  end for }
(12) for  $i = 0$  to  $Len(V_w[]) - 1$  do
(13) 计算  $W_i = decompression(A_0, V_w, 6)$ 
(14) end for

```

2.4 水印相似度计算

为验证水印的提取效果,需计算原始水印 W 与提取到的水印 W_i 间的相似情况。对本文方案,当载有水印信息的数字矢量对象被删除时,属性文件内该对象描述信息仍保留。因此在计算水印相似度时,只考虑内部数据发生改变对 W_i 的影响,此时有

$$\begin{aligned}
 Sim(W, W_i) &= \frac{\sum_{i=1}^n \langle W_i, W_i \rangle / N}{\sqrt{\sum_{i=1}^n W_i^2} \cdot \sqrt{\sum_{i=1}^n W_i^2}}, \langle W_i, W_i \rangle \\
 &= \begin{cases} 1, & W_i = W_i \\ 0, & W_i \neq W_i \end{cases} \quad (1)
 \end{aligned}$$

其中 N 为水印长度。

3 实验与分析

3.1 实验数据及评测标准

信息隐藏方案的测试数据采用标准的哈尔滨市市区数字矢量地图,包括水系、公路交通、区划、服务点等多种图层信息,地图间所含数字矢量对象的数量相差较大。

实验环境基于 MapX 插件的 VC.NET 程序,在保证测试地图等嵌入率的情况下,重点考察算法的误码率(error coding rate, ECR),即提取到的错误水印编码长度(error extraction length, EEL)同水印总长度(total length, TL)的比值:

$$ECR = \frac{EEL}{TL} \times 100\% \quad (2)$$

3.2 精度损失实验

水印信息分散嵌入 MAP 文件的多个对象定义块的冗余空间内,单位块内所嵌入的水印压缩编码长度远远小于冗余空间规模,不会出现数据溢出或被覆盖的情况,对坐标定义块中的地理精度信息无丝毫影响;另一方面,当地理精度发生调整(通常是放大)时,坐标定义块数据更新,对象定义块数据不变,方案的误码率不变。

为了更好地说明问题,本文选择邵承永等人提出的基于统计特征的二维数字矢量地图鲁棒水印算法^[10]进行性能对比(图 2a)。通过应用程序随机生成一幅 80×80 的二值水印图像 W_j (图 2b),经过编码压缩后的水印长度为 1067 比特,这里选用了哈尔滨市服务点数字矢量地图(图 2c),地图比例为 1:480000,对象规模为 5733 个点像素。文献[10]提出

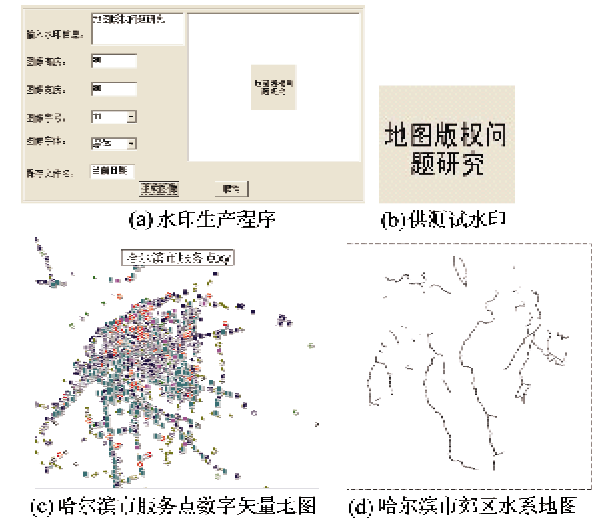


图 2 精度损失实验资料

的方案是通过依赖对象间坐标的相关性在一定程度上实现了地图无损恢复,而哈尔滨市服务点数字矢量地图内的地理标志均为离散的点像素,这使得算法的性能有所降低,地图的精度损失显而易见(表 1)。特别地,随着地图精度不断放大,算法的坐标差值扩大策略逐渐失效,算法的误码率迅速上升(表 2),直至水印提取失败。

表 1 地理精度损失实验随机抽取结果

原始坐标值 (x, y)	水印嵌入后坐标值	
	编码压缩算法	统计特征算法 ^[10]
(193.53627843, 25.34683852)	(193.53627843, 25.34683852)	(193.79351262, 26.521030469)
(256.10257437, 155.36272821)	(256.10257437, 155.36272821)	(256.10285926, 153.43717391)
(96.02520256, 51.36024826)	(96.02520256, 51.36024826)	(97.89103985, 52.05282901)
(321.89352052, 60.25289379)	(321.89352052, 60.25289379)	(321.89352052, 59.01537899)

表 2 地图精度调整实验随机抽取结果

精度调整比例	平均误码率(%)	
	编码压缩算法	统计特征算法 ^[10]
原始精度	0.03	4.68
扩大 10	0.03	12.35
扩大 10^2	0.03	28.79
扩大 10^3	0.03	47.93
扩大 10^4	0.03	62.96

3.3 抗几何攻击实验

对于数字矢量地图信息隐藏方法来说,有效提

高抗几何攻击的性能是公认的研究难点之一。几何攻击方式主要包括对数字矢量地图的剪切、旋转、平移以及投影等。由于几何攻击可同时造成空域及频域范围内的数据扰动,多数信息隐藏算法都会出现性能的大幅降低,甚至失效的情况。

本文选取了剪切、旋转、平移、投影 4 种典型的攻击方式,对算法的抗攻击性能进行评价。我们选择水印 W_0 (图 2b)以及坐标相关性明显的哈尔滨市郊区水系数字矢量地图(图 2d)作为测试资料,数字矢量地图比例为 1:350000,对象规模为 44 个点像素,实验结果如表 3 所示。

表 3 几何攻击实验结果

几何攻击类型		误码率(%)	
		编码压缩 算法	统计特征 算法 ^[10]
剪切	左上 1/6	0.02	2.25
	左侧 1/3	0.03	4.32
	中心 1/6	0.02	3.11
	右下 1/4	0.04	6.28
	左上 1/4		
旋转	顺时针 30 度	0.02	7.89
	逆时针 60 度	0.02	11.23
平移	向右平移 0.2	0.04	59.86
投影	罗宾逊投影	0.03	3.59
	莫卡托投影	0.03	5.59

4 结 论

本文提出了一种容量大、鲁棒性强且具有良好不可感知性能的数字矢量地图信息隐藏方案。该方案通过动态编码表的方式将常规水印转换为长度较短的压缩编码,同时,利用数字矢量地图属性文件中存在大量冗余空间的特点,将水印压缩码嵌入对象定义信息的结束符之后,确保了地图阅读编辑软件无法读取水印信息,而且文件的二进制格式也加大了定位及破译水印编码工作的难度。试验结果表明该方法同已有方法相比,在抗攻击性能和对地理精度的损失方面,可以获得更低的误码率和更优越的应用性能。

参考文献

[1] Ohbuchi R, Ueda H, Endoh S. Robust watermarking of vector digital maps. In: Proceedings of IEEE International Conference on Multimedia and Expo (ICME'02), Lausanne, Switzerland, 2002. 577-580

- [2] Xu D H, Zhu C Q, Wang Q S. A survey of the research on digital watermark for the vector digital map. *Geomatics World*, 2007, 12(6):42-48
- [3] Pan Z H, Li Z J, Gong Z X. A survey of digital watermarking. *Computer & Digital Engineering*, 2008, 36(4): 119-121,133
- [4] Cox I J, Kilian J, Leighton F J, et al. Secure spread specnum watermarking for multimedia. *IEEE Trans on Image Processing*, 1997, 6(12): 1673-1686
- [5] Pu Y C, Du W C, I-Chang Jou. Toward blind robust watermarking of vector maps. In: Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, China, 2006. 930-933
- [6] Wen Q, Sun T F, Wang S X. Concept and application of zero watermark. *ACTA Electronica Sinica*, 2003, 31(2): 214-216
- [7] Kitamura I, Kanai S, Kishinami T. Copyright protection of vector map using digital watermarking method based on discrete Fourier transform. In: Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, 2001, (3): 1191-1193
- [8] Min L Q, Yu Q H. A digital map watermarking algorithm based on discrete cosine transform. *Computer Applications and Software*, 2007, 24(1):147-148,174
- [9] Li Y Y, Xu L P. Vector graphical objects watermarking scheme in wavelet domain. *Acta Photonica Sinica*, 2004, 33(1):97-100
- [10] 邵承永,汪海龙,牛夏牧等. 基于统计特征的二维数字矢量地图鲁棒水印算法. 电子学报,2005. 33(12A): 2312-2316
- [11] Wang X T, Shao C Y, Xu X G, et al. Reversible data-hiding scheme for 2-D vector maps based on difference expansion. *IEEE Transactions on Information Forensics and Security*, 2007,2(3): 311-320
- [12] 邵承永,王孝通,徐晓刚等. 数字矢量地图的无损数据隐藏算法研究. 中国图像图形学报, 2007,12(2): 206-211
- [13] Yin H, Lin C, Qiu F. A survey of digital watermarking. *Journal of Computer Research and Development*, 2005, 42(7):1093-1099
- [14] 钮心忻,杨义先. 基于小波变换的数字水印隐藏与检测算法. 计算机学报,2000,23(1):21-27
- [15] 陈明奇,钮心忻,杨义先. 基于小波变换及数字矢量量化的隐像术. 计算机研究与发展,2001,38(2):199-203
- [16] Huang X, Xu K, Liang Y. Development and application of MapInfo: An example of assessing mineral resources by analyzing multiparameter geological information. *Earth Science Journal of China University of Geosciences*, 2001, 26(2): 189-191
- [17] 齐锐,屈韶琳. 用 MapX 开发地理信息系统. 北京:清华大学出版社,2003

Research on digital watermarking for vector maps based on coding compression

Xia Songzhu, Zhang Jianpei, Yang Jing

(Department of Computer Science and Technology, Harbin Engineering University, Harbin 150001)

Abstract

A novel scheme for hiding vector map information is presented. The scheme uses the method of dynamic coding compression to reduce watermarks to achieve a large watermark capacity, and embeds watermarks into object definition blocks in the attribute file for a vector map. So it is ensured that the watermarks would never be read and displayed when the map is edited or browsed by any tool software because there are many free spaces in the blocks, followed the concrete object definitions. The experimental results show that the scheme has the features of outstanding imperceptibility and zero loss on accuracy, especially, it can strongly resist various geometric attacks such as cropping, rotation and projection, etc.

Key words: vector map, information hiding, coding compression, digital watermarking, lossless