

一种矢量地图无损数字水印技术

孙建国, 张国印, 姚爱红, 武俊鹏

(哈尔滨工程大学计算机科学与技术学院, 哈尔滨 150001)

摘 要: 在保证对数字矢量地图内容无损的前提下, 提出一种基于编码压缩方案的高容量数字水印算法. 根据数字矢量地图的文件组成结构, 将水印信息隐藏于地图结点的属性数据块内. 由于每个属性定义块都保留了冗余空间, 在这些冗余空间内嵌入的水印信息不被任何地图编辑工具阅读或显示, 水印标识因而具有很好的隐蔽性. 实验表明在保持高容量和不可见性的前提下, 算法对地图剪切, 投影等几何变换, 以及格式转换等操作都有很好的抵抗性能.

关键词: 信息隐藏; 数字水印; 矢量地图; 无损

中图分类号: TP309.1 **文献标识码:** A **文章编号:** 0372-2112 (2010) 12-2786-05

Lossless Digital Watermarking Technology for Vector Maps

SUN Jian-guo, ZHANG Guo-yin, YAO Ai-hong, WU Jun-peng

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

Abstract: In ensuring the content of digital vector maps lossless, a high-capacity digital watermarking algorithm is proposed based on encoding scheme. According to the structure of TAB file for digital vector maps, the watermark is hidden into the attribute blocks. As each vector object definition reserves redundant space, these information bits cannot be edited and displayed by any map editing tool, the watermark has good concealment. Experimental results show that the algorithm maintaining high capacity and good invisibility, it also can strongly resist various geometric transformations such as cropping, rotation, projection, etc.

Key words: information hiding; digital watermarking; vector map; lossless

1 引言

数字矢量地图^[1]是一种根据由属性数据定义, 通过拓扑关系约束形成的图形化数字产品. 通常来说, 数字矢量地图的传播和共享主要通过网络或光盘复制, 这就意味着地图所有者经常提供各种各样的副本给使用者, 在此过程中, 地图数据的安全和版权保护都是不可控制的. 为了解决数字矢量地图信息安全问题, 减少数字矢量地图被非法复制而造成的经济损失, 研究者提出了数字矢量地图水印技术和相关方案. 同其它多媒体数据相比, 有关数字矢量地图的数字水印技术研究内容较少, 目前已知的乌拉圭的 The Digital Map Ltd 公司曾针对数字矢量地图成功开发出了著名的数字版权保护软件 MapSN.

同其它领域的数字水印技术的研究内容相似, 数字矢量地图水印技术也可以分为空域和变换域两大主要研究方向. 基于空域的方法简单快捷, 易于实现, 多采用直接修改地理坐标, 调整像素值的方式, 这些技术对地

图内容改变较大, 水印算法的鲁棒性和不可见性均较差, 仅具有研究价值不具备实用性. 基于变换域的水印技术主要包括离散傅里叶变换, 离散余弦变换以及离散小波变换等研究内容^[2]. 同空域方法相比, 变换域方法降低了对地图数据的扰动, 鲁棒性能也有所提高, 但是由于无法满足地图高精度和内容零扰动的实用性标准, 所以仍需要对相关技术进行继续改进.

为了彻底解决上述难题, 德国和中国的研究人员先后提出了可逆数字水印算法^[3]和无损信息隐藏方案^[4]. 前者通过离散余弦变换将地图分成两个区域, 分别定义为隐藏域和搜寻域, 在搜寻域内求得阈值, 借助阈值将水印嵌入到隐藏域内; 而后者通过修改相邻坐标间的差值来嵌入水印信息, 且能够在提取水印的同时使数字地图内容得到无损恢复. 上述两种方法保证了数字矢量地图能够以无损原始态呈现给合法的使用者, 但是对数字矢量地图的内容和类型都有严格要求, 例如对于独立点图层或大比例尺的数字地图, 算法的容量就非常有限; 对于精度调整等常规操作, 算法的抗攻击能力也都较差.

为了提高算法的容量和实用性,本文提出了一种数字矢量地图无损数字水印方案,它能够大容量、无损地向地图内嵌入水印信息,实现一种与数字地图内容无关、格式无关的实用型技术.该算法具有如下特点:(1)可根据用户自定义的编码压缩方案动态调整待嵌入地图内的水印容量;(2)水印信息被嵌入到地图文件的对象定义块内,由于被存储在有效信息的结束符之后,故信息不易察觉,对地图内容和精度更不会产生任何扰动.

2 编码压缩方案

从已发表的有关数字矢量地图水印算法的文献[5]可知:对于数字矢量地图来说,目前可供水印嵌入的信息载体只有地图结点的坐标值和象素值两类.而坐标值和象素值均是地图图元的基本属性,图元与地图内矢量对象的结点数目相当.故一旦地图规模较小(大比例尺地图)或可选结点数目较少(由于高精度问题,关键性结点的坐标值或象素值禁止被调整)时,可嵌入到地图内的水印编码长度必定受限,可重复嵌入的次数更少.

为了解决数字水印算法的容量问题,本文引入了数字水印编码压缩方案.方案包括编码和解码两个过程,分别与数字水印的嵌入过程和提取过程相对应.编码和解码过程都需要根据编码压缩表操作.根据实际需要,由用户自定义所需的定长编码压缩表.

如表 1 所示,为了说明问题,本文定义了 5 比特位的定长编码表.每 5 位二进制的水印编码经过压缩后都可由 1 位压缩字符来表示.字符“ \Rightarrow ”左侧为压缩字符,右侧为原始水印编码.对于 5 位定长编码表来说, $C_5 = \{a, b, \cdots, z, 9, 8, \cdots, 4\}$,码字个数为 32,码率 $R(C_5) = \frac{\log_2 M}{n} = \frac{\log_2 32}{5} = 1$.

表 1 5 位定长编码表

编码 1 区	编码 2 区	编码 3 区	编码 4 区
$a \Rightarrow 00000$	$i \Rightarrow 01000$	$q \Rightarrow 10000$	$y \Rightarrow 11000$
$b \Rightarrow 00001$	$j \Rightarrow 01001$	$r \Rightarrow 10001$	$z \Rightarrow 11001$
$c \Rightarrow 00010$	$k \Rightarrow 01010$	$s \Rightarrow 10010$	$9 \Rightarrow 11010$
$d \Rightarrow 00011$	$l \Rightarrow 01011$	$t \Rightarrow 10011$	$8 \Rightarrow 11011$
$e \Rightarrow 00100$	$m \Rightarrow 01100$	$u \Rightarrow 10100$	$7 \Rightarrow 11100$
$f \Rightarrow 00101$	$n \Rightarrow 01101$	$v \Rightarrow 10101$	$6 \Rightarrow 11101$
$g \Rightarrow 00110$	$o \Rightarrow 01110$	$w \Rightarrow 10110$	$5 \Rightarrow 11110$
$h \Rightarrow 00111$	$p \Rightarrow 01111$	$x \Rightarrow 10111$	$4 \Rightarrow 11111$

在此基础上,如表 2 所示,本文又定义 4 个扩展字符,用以调整水印编码长度.当水印串的长度 $\text{length}(W)$

< 5 时,可结合扩展字符来表示压缩后的水印信息.扩展字符编码表记作 $P = \{0, 2, 3, \#\}$.

表 2 扩展字符编码表

扩展符号	原始编码
0	0
2	00
3	000
#	0000

(1)扩展问题,即水印串的长度 $\text{length}(W) > 5$ 时,可采用在原压缩编码 C_0 的左侧添加扩展字符的方法进行扩展,即 $C_n = \{P\}^* C_5, n > 5$.以此类推,对于表 1 中的 5 位定长编码集合 C_5 来说,6 位定长码集合可表示为 $C_6 = \{0 C_5, 1 C_5\}$,7 位定长码集合可表示为 $C_7 = \{2 C_5, 01 C_5, 10 C_5, 11 C_5\}$.

(2)尾码问题,根据上述扩展规则,可以定义任意长度的编码压缩表.一旦编码表 C_k 确定下来,必然会遇到尾码问题,即水印串的长度 $\text{length}(W)$ 无法被定长 k 整除.设可整除部分的编码为 C_w ,余数部分为 $W_j, W_j = \{W_1, W_2, \cdots, W_j\}, 0 < j < 5$.为此,可在 C_w 之后加入扩展字符来表示剩余水印编码,即 $C_w P_j = \{W_1, W_2, \cdots, W_j, 0_1, 0_2, \cdots, 0_j\}, C_w P_j$ 可由编码表 C_k 中的压缩字符唯一表示.

一般说来,对于给定的数字矢量地图来说,当明确水印编码长度和嵌入强度(重复嵌入的次数)时,就可以制定出具体的编码压缩表.

编码压缩过程发生在水印嵌入操作之前.与编码压缩方案相对应的是解码方案,解码过程发生在水印信息被检测并成功提取之后.为了确定被提取到的水印信息的真实性和完整程度,需要参照事先定义好的编码压缩表,对水印编码进行反向解码,再与原始水印编码进行逐比特位比较,以达到保护数字矢量地图版权的目的.

综上所述,当水印信息具有高容量的特性后,还需要确保嵌入地图内的水印信息不会对地图内容产生扰动,且对于任何地图操作具有很好的适应性和鲁棒性.

3 无损数字水印技术

3.1 数字矢量地图文件结构

从数字矢量地图应用现状看^[6],常见的地图文件类型以 MapInfo 软件支持的 TAB 文件和 ArcGIS 软件支持的 SHP 文件为主,二者之间还可相互转换.

对于 MapInfo 的 TAB 文件来说,主要包括四类系统文件:地理标记文件,结点序列文件,属性定义文件和拓扑约束文件.其中,属性定义文件记录了每一个结点的名称,坐标,资源等信息.在文件中一个结点的完整信息被存储在一个属性数据块内,由于软件版本不同,

所以数据块的大小为 512 字节或 1024 字节两种. 鉴于数字矢量地图产业的迅速发展, 最新版本对数据块又扩展了 200 字节.

在属性定义文件中, 属性数据块包含了对象类型(点图元, 线图元, 面图元)和对象名称两部分内容. 在有效描述信息之后都存储了一个结束符. 结束符之后的数据块均被空置. 经实际验证, 各类数字矢量地图编辑软件均无法读取和显示这些空置区域存储的信息^[7,8].

若将数字水印信息嵌入到这些空置的数据块内, 则地图的精度和内容均不会受到任何扰动, 做到了真正的数据无损. 特别值得注意的是, 当地图结点被删除时, 数据块内的删除标记位会被置为真, 但相关的属性定义信息却仍被存储在数据块内, 故水印信息不会由于结点被删除而被清除. 此外, 当属性信息改变时, 位于数据块末尾的水印信息也不易被覆盖.

综上所述, 将经过编码压缩后的水印信息嵌入到地图属性文件的数据块内, 既保持了高不可见性和大容量, 更对地图内容零扰动.

3.2 数字水印嵌入流程

设数字矢量地图 V 所含的可选结点数目为 N , 数字水印 W 的编码长度为 L . 实验选择了 5 比特压缩编码表, 每个结点数据块内可嵌入 t 位水印编码.

压缩后的数字水印编码会被分散地嵌入到各个结点的属性数据块内, 编码自数据块的末尾开始逆向嵌入, 以加强水印信息的健壮性. 具体嵌入流程如图 1 所示.

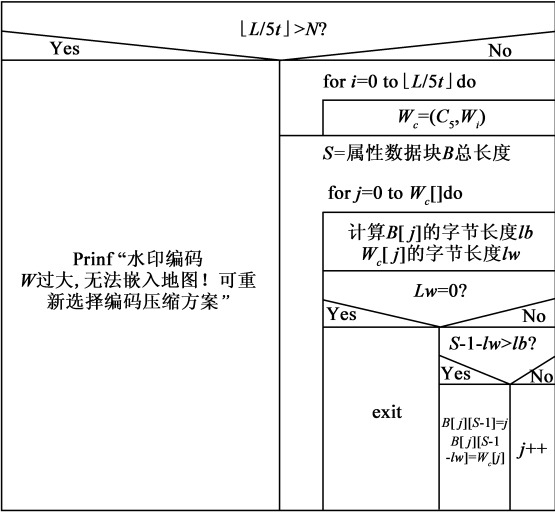


图1 数字水印嵌入流程图

3.3 数字水印提取流程

本文选用的是最新版本的 MapInfo 软件, 属性数据块的大小为 1224 字节. 由数字水印的嵌入过程可知, 含有水印信息的数据块的末位都保留了水印编码的嵌入

顺序, 所以 $int(B[1223]) > 0$. 设数字矢量地图 V 所含的结点总数为 M , 则数字水印技术的提取流程如图 2 所示.

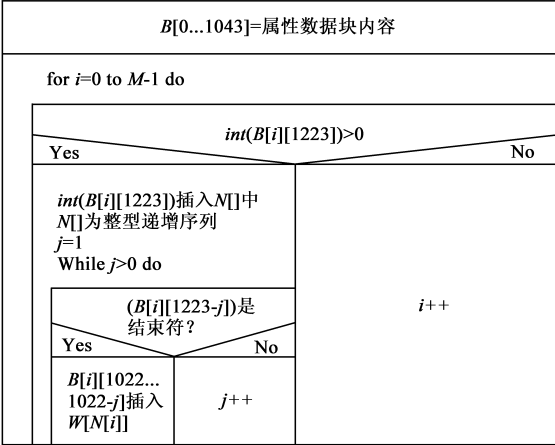


图2 数字水印提取流程图

对于提取到的数字水印编码需要进行解码操作. 根据预先定义好的 5 位定长编码压缩表, 可将水印直接解码为 $WE[]$.

为了验证水印 WE 的正确性和完整程度, 需要计算原始水印 W 与解码水印 WE 的差异情况, 如公式 1 所示, L 为编码总长度, WE_i 与 W_i 分别表示水印编码的一位独立比特值.

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

4 实验分析

4.1 实验数据

本文数字水印算法的测试数据选用实际的哈尔滨市数字矢量地图. 按照城市服务点, 河流水系, 公路铁路交通的分类体系, 数字矢量地图结点数目的分布情况如图 3 所示. 基于 VC6.0 所开发的实验系统如图 4(a)所示, 水印信息如图 4(b)所示.

4.2 无损性分析

随着数字矢量地图应用领域的不断扩大, 对于数字矢量地图水印算法的实用性标准不断提高, 其中首要问题就是必须保证载有水印信息的数字矢量地图其内容是完整且无损的^[9].

如前所述, 压缩后的水印信息被分散地嵌入到多个地图结点的属性数据块中, 这些水印信息与地图有效数据通过结束符彼此隔离, 故不存在对地图内容的扰动. 另一方面, 水印算法对于精度约减、比例尺调整

以及格式转换等常规操作的抵抗性能较好,上述情况下水印信息都不易丢失。

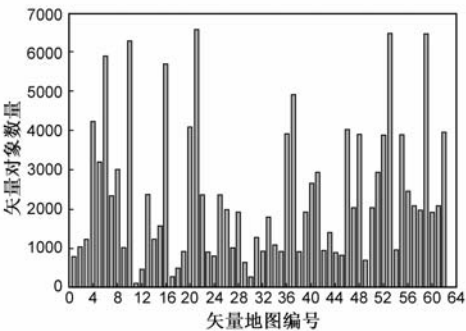
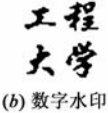


图3 数字矢量地图的结点规模



(a) 数字矢量地图水印系统平台



(b) 数字水印

图4 实验环境

哈尔滨市服务点数字矢量地图(图 5)的比例尺为 1 :480000,共包含 5733 个结点.原始水印编码按照 5 位定长编码压缩规则出来后,其编码长度为 2155 比特.我们选择该幅地图作为算法无损性的实验对象.

地图水印算法的无损性实验主要考虑的内容包括:(1)水印嵌入后,地图结点坐标值和象素值是否被改变;(2)调整比例尺后,地图水印是否能够被完整提取;(3)格式转换后,地图水印是否能够被完整提取.

如表 3 所示,文献[10]将水印嵌入到结点的坐标值中,而本文算法将水印嵌入地图属性文件内,故结点的

象素值不需考虑.

表 3 地理精度无损情况测试

原始坐标(x,y)	水印嵌入后坐标值	
	本文算法	文献[10]
(71.20483911, 293.02384158)	(71.20483911, 293.02384158)	(71.20487103, 293.02389284)
(183.43098254, 123.45878292)	(183.43098254, 123.45878292)	(183.43103293, 123.45879301)
(592.92020145, 18.23498231)	(592.92020145, 18.23498231)	(592.92020982, 18.23498349)
(49.38191941, 345.84920204)	(49.38191941, 345.84920204)	(49.38193452, 345.84923940)
(583.68848576, 413.34523531)	(583.68848576, 413.34523531)	(583.68843491, 413.34523489)

在保持水印等嵌入率的情况下,本文从不同水印算法处理后的数字矢量地图内随机抽取了五个结点坐标值,显而易见,文献[10]所述算法依赖对象间坐标的相关性来实现地图无损恢复,而哈尔滨市服务点矢量地图内均为离散的结点图元,故算法的性能有所降低.

如表 4 所示,当比例尺不断调整,精度不断降低时,文献[10]的坐标差值扩大策略将逐渐失效,误码率也迅速上升,水印完整性被极大破坏.算法的参考指标为误码率,即提取的错误水印编码长度与水印总长度的比值.

表 4 比例尺调整无损情况测试

比例尺 调整比例	误码率(%)	
	本文算法	文献[10]
原始状态	0	0
扩大 10^1	0	2.51
扩大 10^3	0	7.83
扩大 10^5	0	16.39
扩大 10^7	0	28.72

鉴于数字矢量地图的不同用途,使用者对地图精度受损情况的容忍度也不同.从目前的应用前景来看,使用者普遍对地图精度无损要求越来越强烈,已成为地图实用的必备标准之一.

4.3 鲁棒性分析

对于数字矢量地图水印技术来说,除了要满足无损性的必备标准外,还要平衡数字水印算法在不可见性,鲁棒性以及容量三种关键性能间的关系.其中,有效提高水印算法的抗几何攻击性能是公认的研究难点之一,类似的几何攻击本质上是用户对地图所作的各种常规编辑操作.

几何攻击主要包括对数字矢量地图的剪切,旋转,平移以及投影等操作.由于几何攻击容易同时造成地图空域及频域范围内的数据变动,所以很多数字水印算法都会出现性能降低,甚至功能失效的情况.

本文选取剪切、旋转、平移三种常见的操作方式,用

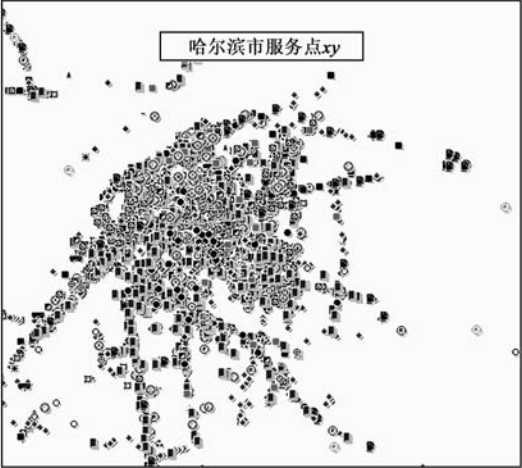


图5 哈尔滨市服务点矢量地图

以测试本文算法的鲁棒性能.数字矢量地图比例尺为 1:1250000,共包含 2193 个结点,实验结果如表 5 所示.

表 5 鲁棒性能测试

几何编辑操作		误码率(%)	
		本文算法	文献[10]
剪切	左上 1/4	0.35	2.13
	右侧 1/2	1.12	4.47
	中心 1/4	2.08	1.92
	左上 1/8 同时右下 1/8	0.74	6.10
旋转	顺时针 90 度	0	27.91
	逆时针 145 度	0	21.38
平移	左平移 1mm	0	59.75
	右平移 2mm	0	41.53

4.4 不可见性分析

经实际使用发现,运用 MapInfo 等地图编辑软件对地图任意结点的属性数据进行查看或编辑时,均无法显示或改动嵌入数据块内的水印压缩信息.此外,当地图结点被删除时,通过程序仍可读取结点数据块内的二进制水印编码信息.

由上述实验结果可知,本文提出的数字水印算法具有很好的不可见性.

5 结论

本文提出了一种大容量、强鲁棒性且具有高不可见性的数字矢量地图水印算法.算法通过选定的定长编码表将水印编码压缩为长度较短的字符,用以解决水印算法的容量问题;同时,利用数字矢量地图结点的属性数据块保留大量冗余空间的特点,将水印压缩码分散嵌入到数字矢量地图各结点的数据块内;水印信息被存储在属性信息结束符之后,可确保地图编辑软件无法读取.

同已有数字水印算法相比,本文算法在容量,无损性以及抗攻击性方面都有较大突破,具备了更好的实用性能.

参考文献:

[1] Xu D H, Zhu C Q, Wang Q S. A survey of the research on digital watermark for the vector digital map [J]. Geomatics World, 2007, 12(6): 42 – 48.

[2] 温泉,孙铤锋,王树勋.零水印的概念与应用[J].电子学报, 2003, 31(2): 214 – 216.

Wen Q, Sun T F, Wang S X. Concept and application of zero watermark[J]. Acta Electronica Sinica, 2003, 31(2): 214 –

216. (in Chinese)

[3] Voigt M, Yang B, Busch C. High-capacity reversible watermarking for 2D-vector data[A]. Security, Steganography, and Watermarking of Multimedia Contents VII[C]. San Jose, United States: SPIE, 2005. 409 – 417.

[4] Pu Y C, Du W C, I C J. Toward blind robust watermarking of vector maps [C]. 18th International Conference on Pattern Recognition[C]. Hong Kong, China: IEEE, 2006. 930 – 933.

[5] Cox I J, Kilian J, Leighton F J, Shamoon T. Secure spread specnum watermarking for multimedia[J]. IEEE Trans on Image Processing, 1997, 6(12): 1673 – 1686.

[6] 郭萌,张鸿宾,魏磊.二值图像中的数据隐藏算法[J].电子学报, 2009, 37(11): 2409 – 2415.

Guo M, Zhang H B, Wei L. Data hiding in binary images[J]. Acta Electronica Sinica, 2009, 37(11): 2409 – 2415. (in Chinese)

[7] 秦拯,易叶青,林亚平.基于 JADE 算法的鲁棒性数字水印[J].电子学报, 2008, 36(6): 1149 – 1153.

Qin Z, Yi Y Q, Lin Y P. Robust watermark based on JADE algorithm[J]. Acta Electronica Sinica, 2008, 36(6): 1149 – 1153. (in Chinese)

[8] 段贵多,李建平,廖建明.一种新颖的用于图像内容认证、定位和恢复的半脆弱数字水印算法研究[J].电子学报, 2010, 38(4): 842 – 847.

Duan G D, Zhao X, Li J P, et al. A novel semi-fragile digital watermarking algorithm for image content authentication, localization and recovery[J]. Acta Electronica Sinica, 2010, 38(4): 842 – 847. (in Chinese)

[9] Pan Zhi-hong, Li Zhi-juan, Gong Zi-xia. A Survey of digital watermarking[J]. Computer & Digital Engineering, 2008, 36(4): 119 – 121, 133.

[10] 邵承永,汪海龙,牛夏牧等.基于统计特征的二维矢量地图鲁棒水印算法[J].电子学报, 2005. 33(12A): 2312 – 2316.

Shao C Y, Wang H L, Niu X M, et al. A robust watermarking algorithm for 2D vector maps based on statistic detection[J]. Acta Electronica Sinica, 2005. 33(12A): 2312 – 2316. (in Chinese)

作者简介:

孙建国 男,1981 年出生,黑龙江省巴彦县人,博士,哈尔滨工程大学讲师.主要研究方向为信息安全、嵌入式系统设计.

E-mail: sunjianguo@hrbeu.edu.cn