

# 基于坐标映射的矢量地理数据全盲水印算法

杨成松<sup>1)</sup> 朱长青<sup>2)</sup> 陶大欣<sup>1)</sup>

<sup>1)</sup>(信息工程大学测绘学院, 郑州 450052) <sup>2)</sup>(南京师范大学虚拟地理环境教育部重点实验室, 南京 210097)

**摘要** 通过对全盲水印特点的分析,提出了将文字按照字符编码方式直接转化为二进制位的水印信息生成方式。基于矢量地理数据的空间定位特性,研究了水印嵌入与提取中坐标映射机制和映射函数构造原则。利用提出的水印生成和坐标映射方法,建立了一种矢量地理空间数据全盲水印算法。实验结果表明,提出的算法能抵抗矢量数据中常见的数据压缩、增点、删点、编辑、裁剪以及平移等攻击,具有很好的鲁棒性。

**关键词** 数字水印 矢量数据 全盲水印 映射

中图法分类号: TP301.6 文献标志码: A 文章编号: 1006-8961(2010)04-684-05

## A Blind Watermarking Algorithm for Vector Geo-spatial Data Based on Coordinate Mapping

YANG Chengsong<sup>1)</sup>, ZHU Changqing<sup>2)</sup>, TAO Daxin<sup>1)</sup>

<sup>1)</sup>(Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052)

<sup>2)</sup>(Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing 210097)

**Abstract** By the analysis of the properties of blind watermarking, a method for generating watermark information which transforms characters directly into binary code is proposed. Secondly, the mechanism of coordinate mapping and the principles of constructing mapping function for watermarking embedding and extracting are studied based on the location characteristic of vector geo-spatial data. Then, a blind watermarking algorithm for vector geo-spatial data is obtained using the proposed approaches of generating watermark information and coordinate mapping. The experiments conducted in this paper show that the algorithm is with good robust and can resist some usual attacks such as data compressing, adding, deleting, editing, clipping, translating and so on.

**Keywords** digital watermarking, vector geo-spatial data, blind watermarking, mapping

## 0 引言

数字水印技术是随着数字化的进程而发展起来的信息安全的前沿技术,它通过在源数据中嵌入秘密水印信息,将水印与源数据紧密结合并隐藏其中,成为源数据不可分离的一部分,由此来确定版权拥有者、所有权认证、跟踪侵权行为等。数字水印技术在图像、音频、视频以及文字等数字产品的安全保护

方面得到了广泛地应用<sup>[1-2]</sup>。近年来,随着地理空间数据安全保护需求的日益增加,数字水印技术在矢量地理数据安全保护方面也得到了重要的应用<sup>[3]</sup>。

数字水印按照水印检测时是否需要原始数据与密钥等其他附加信息参与可以分为非盲水印、半盲水印和全盲水印3种。非盲水印检测时需要原始数据的参与,半盲水印检测时不需要原始数据的参与,但需要密钥或者其他附属信息的参与,全盲水印检

基金项目:国家高技术研究发展计划(863)项目(2006AA12Z223, 2009AA12Z228)

收稿日期:2010-01-14;改回日期:2010-01-21

第一作者简介:杨成松(1982—),男,苗族。郑州信息工程大学测绘学院博士研究生。研究方向为空间数据处理、矢量地理数字水印。

E-mail: ycsdongshang@163.com

通讯作者简介:朱长青, E-mail: zcq88@263.net

测时不需要原始数据也不需要密钥及其他附加信息的参与。从目前已有关于矢量数据数字水印技术文献来看,闵连权提出了“分类栅格化”(映射)的水印嵌入位置的确定方法,实现了盲水印,取得了较好的效果,但对于“分类栅格化”(映射)的原则没有进行阐述<sup>[4]</sup>,同时由于采用的是文字栅格化的方式生成水印信息,嵌入的水印长度可能会随着原始字符长度的变化而变化。由于不能抵抗简单的同步攻击,如删点、坐标倒序等常用攻击很难真正意义上实现全盲水印。其他大部分算法都属于非盲检测和半盲检测<sup>[5-9]</sup>,水印检测都需要原始数据或者一些附加信息的参与,这影响了水印算法的实用性。目前迫切需要全盲水印算法来满足实际应用中矢量地理数据版权保护的需求。

考虑到实际应用的需求,基于矢量地理数据自身特性,本文对矢量地理数据全盲水印技术进行了研究。

## 1 矢量地理数据特性分析

矢量地理数据作为地理空间数据的重要组成部分,具有便于面向现象的数据表示、数据结构紧凑、冗余度低、有利于网络分析、图形显示质量好、精度高等主要特点<sup>[10-11]</sup>。对于矢量地理数据的使用,大部分的原因是用户追求其地理数据精度的权威性,因此,对于矢量数据的任何更改都是非常谨慎的。特别地,对地理数据直接进行平移、放大(缩小)、旋转从而影响数据精度是非常危险的,也是不允许的<sup>[12]</sup>。

当然,在不影响数据精度和质量的前提下,对矢量数据的一些处理是被允许的,常见的数据处理方式有数据压缩、数据删点、数据增点、数据点的移位、数据裁剪等。同时,和栅格数据不同,在不影响拓扑关系的前提下,坐标文件中点、线、面状要素之间的先后关系是可变的。这些决定了不能根据嵌入水印信息的先后顺序来检测水印信息。常见的增(删)点、坐标顺序打乱等操作决定了不能直接把图像中的盲水印算法移植到矢量地理数据中来。

作为地理空间数据的重要组成部分,与一般的栅格图数据相比,矢量地理数据具有空间性这一特点,即矢量地理空间数据具有精确的空间定位特性,数据点的增删以及数据点之间先后顺序的改变不会影响任何数据点的空间定位特性。

矢量地理空间数据具有的空间定位特性是矢量地理空间数据全盲水印研究的基础。

## 2 矢量地理数据全盲水印算法

### 2.1 水印的生成

无意义水印信息检测时一般需要密钥的参与,因此,考虑到要实现全盲检测以及实际应用中批量检测的需要,嵌入的水印信息为有意义的文字。

有意义的文字转换成水印信息通常是先栅格化显示字符,然后对字符按行列进行扫描,得到水印信息。这种方式的转换是不可逆的,即栅格化得到的水印信息很难转换回有意义的字符。同时,栅格化得到的水印信息一般比较长,水印长度随字符长度的变化而变化,难以实现真正意义上的盲水印和批量检测,因为每次嵌入的水印信息长度不是固定的。

为了在检测过程中实现批量的盲水印检测,提出一种直接把文字字符按照文字编码方式转化为二进制位的水印信息生成方式。这种方式的转换方式是可逆的,并且使用较短的水印长度就能嵌入较长的水印信息。

由于一个水印位的改变可能会影响水印检测结果,应该保证每个位的水印都嵌入到了水印中,因此水印信息长度不宜过长。考虑到每次嵌入的字符长度可能不一样,为了实现全盲检测,嵌入的水印长度设为固定的长度,比如说设定最大长度为240,能最大嵌入15个汉字或者30个英文字母,这在大多数情况下能满足实际应用的需求。当嵌入的字符经过编码转换后达不到固定长度时,一般采用扩容的方法用固定的位序列对二进制位进行扩容。这样每次嵌入的水印长度是一致的,有利于盲检测的实现。

二进制位扩容之后,为了增强算法的鲁棒性,对扩容后的二进制位进行置乱处理。置乱处理后得到待嵌入的水印信息。水印信息生成流程如图1所示。

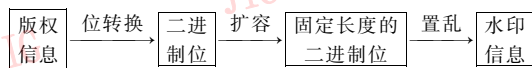


图1 水印信息生成

Fig. 1 Watermark generation

### 2.2 水印的嵌入

水印的嵌入需要解决的关键问题主要有3个:一是水印信息嵌入位置的确定,即每个水印位对应的数据点;二是水印信息的嵌入方式;三是嵌入强度

的确定。其中前两个问题的解决直接关系到水印算法是否为盲检测算法。

在水印嵌入过程中,首先需要解决的就是水印信息的嵌入位置问题,即每个水印位对应的数据点。目前,大部分算法一般都是按照坐标在文件中出现的先后位置依次嵌入水印信息,这很难做到盲检测,因为当嵌入水印信息的坐标先后顺序发生变化和遇到删(增)点攻击时,根据坐标点的先后顺序来提取水印信息将变得非常困难。

常用的算法之所以不能实现盲检测,从本质上来讲是因为坐标点在文件中出现的顺序是相对的、易变的。通过矢量地理空间数据特性分析可知,矢量数据具有精确的空间定位性,坐标点先后顺序发生变化和遇到删(增)点攻击时坐标值是不会发生变化的。坐标点具有的空间定位特性为水印算法实现盲检测提供了可能。

设水印生成的固定长度为  $N$ ,采用坐标映射的思想<sup>[4]</sup>,可以把坐标文件中所有的坐标点按照其坐标值均匀地映射到  $[0, N - 1]$  上,令映射函数为  $f(x, y)$ ,水印信息为  $\mathbf{W} = \{w[i]\} (0 \leq i \leq N - 1)$ ,即要求  $0 \leq f(x, y) \leq N - 1$ 。同时,考虑到算法的鲁棒性, $f(x, y)$  应尽量均匀地把  $(x, y)$  映射到  $[0, N - 1]$  上。

找到了嵌入位置之后,嵌入方法则可以表示为

$$x \oplus w[f(x, y)] \quad (1)$$

$$y \oplus w[f(x, y)] \quad (2)$$

其中,式(1)表示在  $x$  坐标中嵌入水印,式(2)表示在  $y$  坐标中嵌入水印, $\oplus$ 表示水印嵌入方式。

映射函数  $f(x, y)$  是多对一的映射,即通过映射关系, $i(0 \leq i \leq N - 1)$  可能对应多个坐标点,也就是说在嵌入过程中,同一个水印位将可能被嵌入到多个坐标点中,这使得算法具有抗裁剪(裁减)、增删点的能力。

映射函数  $f(x, y)$  的确定是嵌入算法中最为关键的部分, $f(x, y)$  确定的基本原则为  $f(x, y)$  能尽量均匀地映射到  $[0, N - 1]$  上,同时, $(x, y)$  的细微变化不能显著影响  $i(0 \leq i \leq N - 1)$  对映射函数  $f(x, y)$  所对应的坐标点  $(x, y)$  的个数。

当映射函数  $f(x, y)$  满足上述原则之后,在数据处理过程中,点位坐标  $(x, y)$  一般不会发生大的变化,在常见的删点、裁剪攻击下, $f(x, y)$  不会发生变化,因此能有效地抵抗同步攻击,有利于盲检测的实现。

当然,和其他水印算法一样,数据点不能太少,当数据点少于  $N$  时将无法嵌入水印信息。实验表明,当数据点个数大于  $3N$  时,效果比较理想。

在水印嵌入位置确定之后,需要解决的就是水印信息嵌入方式的选择,最低有效位嵌入、基于比较的嵌入、基于量化的嵌入是实现盲检测的理想嵌入方式,考虑到算法的抗攻击能力,采用了基于量化的水印嵌入方法。

### 2.3 水印的提取和恢复

水印的提取其实是水印嵌入的逆过程,因为每个水印位  $(w[i])$  可能进行了多次的嵌入,因此,在水印提取时,需要做一些特殊处理。

定义一个与水印序列等长的标志序列  $\mathbf{B} = \{b[i]\} (0 \leq i \leq N - 1)$ ,初始化  $b[i]$ ,令  $b[i] = 0$ 。对于所有的  $(x, y)$  坐标点,令

$$b[f(x, y)] += 2D(x) - 1 \quad (3)$$

$$b[f(x, y)] += 2D(y) - 1 \quad (4)$$

式中, $D(\cdot)$  表示水印信息的提取法则, $D(\cdot) \in \{0, 1\}$ ,因此  $(2D(\cdot) - 1) \in \{-1, 1\}$ ;  $D(x)$  表示在  $x$  中提取水印信息, $D(y)$  表示在  $y$  中提取水印信息。因次,当  $b[i]$  不为 0 时,水印提取结果如式(5)所示。

$$w[i] = \frac{b[i]}{|b[i]|} + 1 \quad (5)$$

当  $b[i]$  等于 0 时,表示  $b[i]$  中检测不到水印信息。

提取之后的水印信息是无意义的 0,1 二进制位,需要将其恢复为有意义的版权信息,这一过程其实就是水印信息生成的逆过程。首先对水印信息进行逆置乱,然后根据固定的扩容位来得到嵌入字符的长度,最后根据嵌入字符的长度即可以把二进制位恢复成有意义的字符。

## 3 实验分析

运用提出的水印算法对一幅等高线地图数据(如图 2 所示,160 144 个坐标点)进行实验分析,字符“版权保护”作为水印信息。

### 3.1 可视化比较

对嵌入水印前后的数据进行可视化比较,图 3 是可视化比较后的局部放大图,其中实线为原始数据,虚线为嵌入水印后的数据。从可视化比较可以

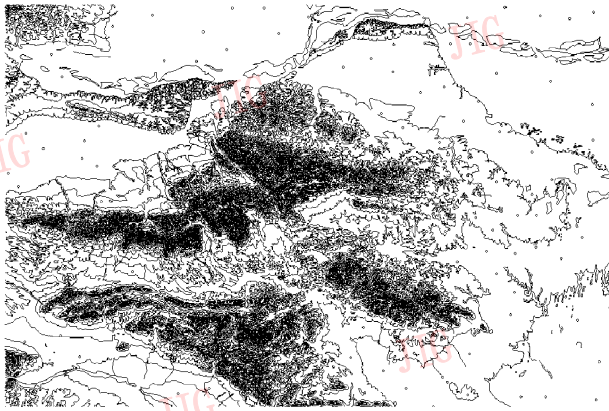


图 2 原始等高线数据

Fig. 2 Original contour line data

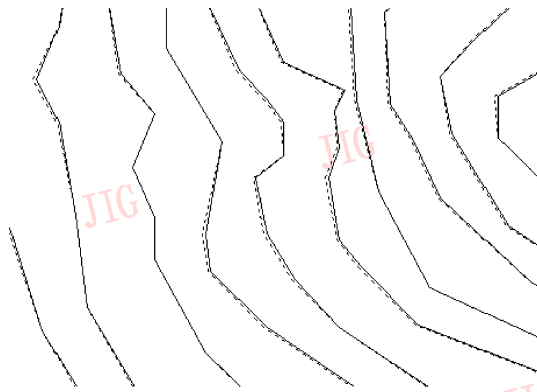


图 3 可视化比较

Fig. 3 The comparing of visualization

看出,水印的嵌入不影响数据的可视化质量。

### 3.2 误差分析

数据精度一直是数据生产者 and 使用者非常关心的问题,水印嵌入对数据精度的影响应该控制在一定范围之内,即水印的嵌入不能影响矢量数据的空间定位特性。对嵌入水印前后的数据进行误差比较可以看出,水印的嵌入不会影响数据的精度,比较结果如表 1 所示。

表 1 嵌入误差分析

Tab. 1 The error analysis

误差大小	数据点个数	所占百分比/%
0	70 454	43.994 156
1	22 608	14.117 294
2	67 082	41.888 550
>2	0	0

### 3.3 鲁棒性分析

对嵌入水印的等高线数据文件进行不同程度的

数据压缩、增(删)点、裁剪、平移、噪声(数据编辑)等攻击,检测结果如表 2 所示,对应检测结果的√表示检测出水印信息,×表示检测不出水印信息。

表 2 攻击实验

Tab. 2 Attacking experiments

攻击方式	攻击强度/检测结果					
数据压缩	压缩百分比	11.8	22.1	31	38.4	44.2
	检测结果	√	√	√	√	√
数据增点	增点百分比	10	20	30	40	50
	检测结果	√	√	√	√	√
数据删点	删点百分比	10	20	30	40	50
	检测结果	√	√	√	√	√
噪声攻击	攻击百分比	10	20	30	40	50
	检测结果	√	√	√	√	√
数据裁剪	裁剪百分比	84	68.8	56	47.5	20
	检测结果	√	√	√	√	√
数据平移	平移量	3	5	7	9	11
	检测结果	√	√	√	×	×

表 2 中压缩百分比指运用道格拉斯-皮克压缩算法对数据进行压缩的压缩比;噪声攻击是在一定量的数据上叠加服从 $[-5, 5]$ 均匀分布的噪声;裁剪攻击是把嵌入水印后的数据裁剪一部分下来进行水印检测,裁剪百分比是指裁剪下来的数据和原始数据的比值;数据平移攻击是对数据同时进行横、纵坐标方向的平移。

从表 2 可以看出,算法能抵抗矢量地理数据数据处理中常用的数据压缩、增点、删点、编辑、裁剪,同时也能抵抗轻微的数据平移攻击,具有较好的鲁棒性。

## 4 结 论

基于矢量地理数据的空间定位特性,对矢量地理空间数据数字水印算法进行了研究,通过对水印生成算法以及坐标映射原则的研究,提出了一种基于坐标映射的矢量地理空间数据全盲水印算法,并对算法的可用性与鲁棒性进行了实验。实验表明,提出的算法能抵抗矢量数据中常见的数据压缩、增点、删点、编辑、裁剪和平移攻击。同时,本算法属于全盲检测算法,检测过程中不需要任何附加信息的参与,算法具有较好的实用性,对于矢量地理数据水印算法的研究和应用具有一定的指导作用。

## 参考文献 (References)

- [1] Sun Shenghe, Lu Zheming, Niu Xiamu, et al. Digital Watermarking Technique and Application[M]. Beijing: Science Press, 2004: 10-14. [孙圣和, 陆哲明, 牛夏牧, 等. 数字水印技术及应用[M]. 北京: 科学出版社, 2004: 10-14.]
- [2] Zhong Hua, Zhang Xiaohua, Jiao Licheng. Digital Watermarking and Image Authentication—Algorithm and Application[M]. Xi dian University Press, 2006: 1-2. [钟桦, 张小华, 焦李成. 数字水印与图像认证——算法及应用[M]. 西安: 西安电子科技大学出版社, 2006: 1-2.]
- [3] Li Yuanyuan, Xu Luping. A blind watermarking of vector graphics images[C]//Proceedings of the 5th International Conference on Computational Intelligence and Multimedia Applications(ICCIMA' 03). Washington DC, USA: IEEE Computer Society Press, 2003: 424.
- [4] Min Lianquan. A robust digital watermarking in cartographic data in vector format[J]. Acta Geodaetica et Cartographica Sinica, 2008, 37(2): 262-267. [闵连权. 一种鲁棒的矢量地图数据的数字水印[J]. 测绘学报, 2008, 37(2): 262-267.]
- [5] Nikolaidis N, Pitas I, Solachidis V. Fourier descriptors watermarking of vector graphics images [C]//Proceedings of International Conference of Image Processing. Washington, DC, USA: IEEE Computer Society Press, 2000, 3: 10-13.
- [6] Solachidis V, Nikolaidis N, Pitas I. Watermarking polygonal lines using fourier descriptors [J]. IEEE Computer Graphics and Applications, 2004, 24(3): 44-51.
- [7] Voigt M, Yang B, Busch C. Reversible watermarking of 2D-vector data[C]//Proceedings of 2004 Multimedia and Security Workshop on Multimedia and Security. New York, NY, USA: ACM Press, 2004: 160-165.
- [8] Zhu Changqing, Yang Chengsong, Li Zhongyuan. An anti-compression watermarking algorithm foe vector map data [J]. Zhengzhou Institute of Surveying and Mapping, 2006, 23(4): 281-283. [朱长青, 杨成松, 李中原. 一种抗数据压缩的矢量地图数据数字水印算法[J]. 测绘科学技术学报, 2006, 23(4): 281-283.]
- [9] Zhu C Q, Yang C S. A watermarking algorithm for vector spatial geo-data based on integer wavelet transform [C]//Proceeding of the XXI International Congress for Photogrammetry and Remote Sensing, Boca Raton, FL, USA: CRC Press, 2008: 15-18.
- [10] Guo Rengzhong. Spatial Analysis[M]. Beijing: High Education Press, 2001: 13-17. [郭仁忠. 空间分析[M]. 北京: 高等教育出版社, 2001: 13-17.]
- [11] Hua Yixin, Wu Sheng, Zhao Junxi. The Principle and Technique of Geographic Information System [M]. Beijing: PLA Press, 2001: 165-166. [华一新, 吴升, 赵军喜. 地理信息系统原理与技术[M]. 北京: 解放军出版社, 2001: 165-166.]
- [12] Jia Peihong, Ma Jingsong, Shi Liangzhao, et al. Technical methods for encrypting and hiding digital watermark in GIS spatial data [J]. Geomatics and Information Science of Wuhan University, 2004, 29(8): 747-750. [贾培宏, 马劲松, 史照良, 等. GIS空间数据水印信息隐藏与加密技术方法研究[J]. 武汉大学学报(信息科学版), 2004, 29(8): 747-750.]