

doi:10.3772/j.issn.1002-0470.2014.12.004

基于内容扰动的二维电子海图可逆水印技术研究^①

赖明珠^{②*} 孙建国^{③*} 张国印^{*}

(* 哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

(** 哈尔滨理工大学软件学院 哈尔滨 150080)

摘 要 针对二维电子海图对精确度和内容完整性要求高,常规的水印嵌入策略经常会引起电子海图内容的扰动,较难实现实用目标的问题,研究了一种二维电子海图可逆水印技术。一方面通过对海图内容特征的研究,获得了一个内容扰动控制模型;另一方面为了进一步提高海图水印的实用性,提出了基于海图数据组织结构和空间特征的可逆水印方案。实验和研究表明,与同类方法相比,该技术能够适应海图常规操作所带来的数据扰动,在应对图元删除、曲线拟合等方面具有较好的鲁棒性。

关键词 电子海图,信息隐藏,内容扰动,数字水印,可逆

0 引 言

二维电子海图是一种最重要的数字海图,广泛应用于海军装备、海洋开发、海洋工程领域,除了具有高保真和高精度外,它还具有数据耦合度低、图元类型丰富的特点。目前能够查阅到的针对数字海图水印方面的文献较少,文献[1]提出的用于海图版权标识的小波域数字水印技术是检索到的为数不多的电子海图水印技术之一。造成这种状况的原因是电子海图的数据结构复杂,且应用范围较为单一,故研究人员较少;另一个方面,纵观国内外,由于电子海图更多具有军事用途的特点,电子海图的制作和管理都受到严格控制。

当前电子海图的应用趋势日益明显,安全需求

日益增强,而二维电子海图安全性的研究进展却相对缓慢。根据安全水印实现方法的不同,已有二维电子海图安全水印或相近算法大致可分为两类,即基于结构性复合的安全水印和基于内容复合的安全水印。本课题根据国内外相关研究的现状,主要从水印的复合功能角度进行探索,在充分考虑电子海图的数据特性和保持水印方法在水印容量、鲁棒性等方面的稳定性的基础上,提出了一种基于内容扰动模型的二维电子海图可逆水印方案,以期数字海图网络化安全管理、可靠性应用提供一种新的技术。

1 基于特征的扰动模型

为了获得一个理想的水印载体集合,能够隐蔽

① 国家自然科学基金(61202455),教育部博士点基金(20112304120025),黑龙江省自然科学基金(F201212)和教育部中央高校基础研究基金(HEUCF100612)资助项目。

② 女,1973年生,讲师;研究方向:信息安全,数字水印。E-mail:laimingzhu@hrbeu.edu.cn

③ 通讯作者,E-mail:sunjianguo@hrbeu.edu.cn

(收稿日期:2014-06-26)

地、高可靠地嵌入足够容量的水印信息,首要问题是选择并优化可供水印嵌入的数据载体集合;另一方面需要制定向载体内嵌入水印编码的策略,即获得较好的鲁棒性,且需要密切结合海图数据的多种静态结构特性和信息属性。初步分析表明,地图的比例尺缩放、精度调整、图元变更、分层显示等操作对上述特性没有影响。因此,可根据载体数据的特性与这些操作的行为学原理,确定二维电子海图水印高鲁棒性的表达方法。

1.1 海图特征概述

如图 1 所示,根据二维数字海图的基本特征,构建内容特征模型,该模型包括属性信息、几何信息以及关系信息三个重要组成部分。基于该特征模型,海图的数据组织结构可被表示为 $V = \{Da,Dg,Dc\}$, Da 、 Dg 、 Dc 分别代指属性、几何和关系特征。

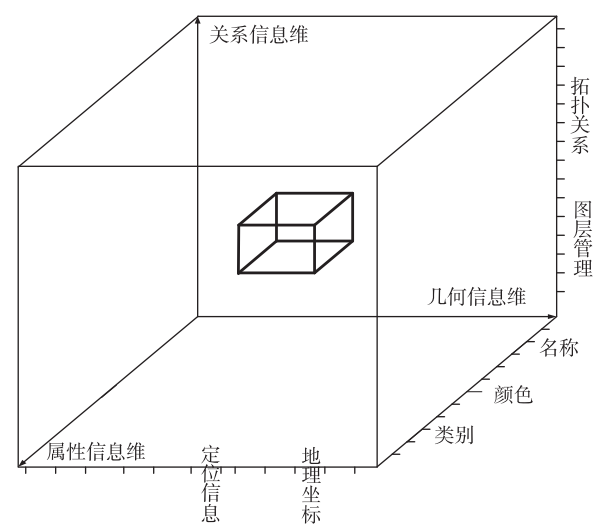


图 1 二维电子海图内容特征模型

进一步,每一类特征包含的具体内容被称为要素,具体表示为特征所属的一组 n 元序列,即 $D_x = \langle e_{n-1}, \cdots, e_1, e_0 \rangle$ 。该序列是有序的,根据要素稳定性由强到弱排列。“ e_i 是稳定的”是指将水印信息嵌入 e_i 后具有较好的鲁棒性。

对于研究数字水印方法来说,海图的内容特征模型还可以进一步简化为属性信息、空间结构和附属信息。

对于海图数字水印研究,附属信息具有更加丰富的特征,可以灵活地嵌入水印信息。附属信息的要素较多,要素之间彼此独立。附属信息可表示为

$$Ds = \begin{bmatrix} e_i & \cdots & e_k & \cdots & e_j \\ r & \cdots & s & \cdots & t \end{bmatrix}$$
, 要素 $e_i, \cdots, e_k, e_j \in Ds$, 且 $\{e_i \cup \cdots \cup e_k \cup e_j\} = Ds$, $\sum (r, \cdots, s, t) = 1$, 稳定参数 $r, \cdots, s, t \in (0, 1)$ 。水印的嵌入对象和容量受稳定参数制约。

简化模型的另一个特点就是空间信息特征相对单一,包括图层、位置、坐标、拓扑结构四类核心内容。空间信息维是数字矢量地图的关键特征,版权标识的嵌入主体应集中于维度下的各种要素。

1.2 数据扰动模型

作为一种重要的地理信息工具,二维电子海图的数据精度通常要保留到 10^{-6} 到 10^{-8} 之间,为了确保二维电子海图的数据无损要求,需要实时调控水印嵌入和提取等环节所可能引入的误差,并制定误差的控制策略。

为此,建立了海图数据内容扰动模型(图 2),分析水印嵌入强度对海图原始数据精度的影响,控制

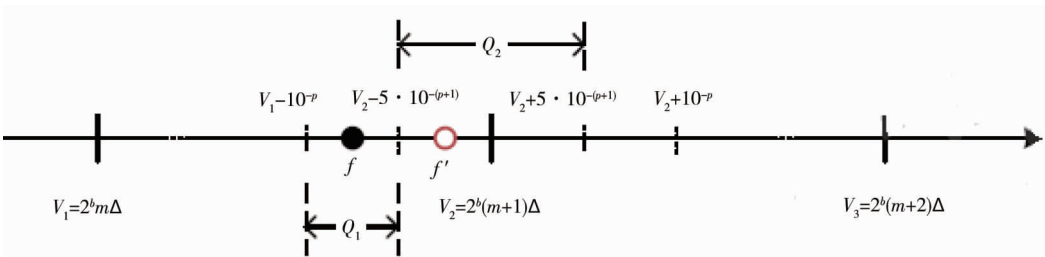


图 2 电子海图内容扰动模型

好误差,然后通过理论分析与实验验证,选择恰当的水印容量和嵌入参数,设误差系数为 f ,地图原始精度为 p ,水印嵌入强度为 b ,单位嵌入系数值为 Δ ,则该地图的扰动区间 $m = \left\lfloor \frac{f}{2^b \times \Delta} \right\rfloor$,即若要达到数据无损的目标,必须确保水印嵌入系数值处于该区间内,条件为:

$$\begin{cases} \Delta > 5 \cdot 10^{-(p+1)} \\ 2^b \times 5 \cdot 10^{-(p+1)} \leq f - m \times 2^b \times \Delta \\ \leq 2^b \times \Delta - 2^b \times 5 \cdot 10^{-(p+1)} \end{cases} \quad (1)$$

在此调整范围,可确保海图数据精度在 10^{-6} 到 10^{-8} 之间不会被改变,在具体实现时,可将地理坐标信息转换为整数值,即扩大为原值的 10^8 到 10^{10} 之间,然后进行整数域的数据变换。

在具体实施时,我们将按照扰动区间规模以及可嵌入水印对象个数,将数字水印编码逐比特嵌入到每个对象的定义块中。具体方案将在算法描述中详细介绍。

2 可逆水印方案

2.1 海图文件的组织结构

如前所述,二维电子海图由属性信息、空间信息和附属信息三部分组成,这些内容以数据文件的方式被存储^[2,3]。以主流的 ArcGIS 应用环境为例,海图文件的核心组织结构为 SHP 文件。SHP 文件是标准的结构化组织模式,包括四类实体,如表 1 所示。

(1)结构说明文件 TAB:定义电子海图的图层结构,包括基本图层、色彩渲染图层等。

(2)属性描述文件 DAT:用来描述海图每一个地物对象的属性信息。

(3)对象索引文件 ID:实现地物对象与其属性信息的一一对应,又称为交叉索引文件。

(4)空间描述文件 MAP:记录地物对象的空间

特征,如几何数据以及关系数据等。

表 1 电子海图文件结构说明

序号	名称	说明
1	TAB 文件	变量定义
2	DAT 文件	资源索引描述信息
3	ID 文件	对象索引描述信息
4	MAP 文件	对象数据定义信息

当水印信息嵌入到海图中后,通常会伴随结点存储到海图文件的属性描述或空间描述文件中。

属性描述和空间特征文件通常由大小为 512 或 1024 字节的数据块组成。如表 2 所示,数据块由文件头、索引块、定义块和资源块组成。

表 2 数据块结构

块序号	名称	说明
0	HEADER BLOCK	文件头块
1	INDEX BLOCK	索引块
2	OBJECT DEFINITION BLOCK	对象定义块
3	COORDINATE DEFINITION BLOCK	坐标定义块
4	DELETED BLOCK	删除块
5	RESOURCE BLOCK	资源块

文件头的块标记为 0,长度固定,用于描述数据块内各子块所占字节长度,以及每个子块的起始位置等基础信息。

索引块的块标记为 1,用于该数据块所对应的地物标记。通过 ID 文件,在 TAB 文件中快速查找空间内的对象。

对象定义块的块标记为 2,记录了每个对象的名称和位置信息等。

坐标定义块的块标记为 3,通常由子对象和实际空间坐标组成,而孤立的空间对象仅由坐标组成。坐标定义块的冗余空间较大。

删除块的块标记为 4,当某对象被删除时,该对

象的删除标记发生改变,但对对象的其他信息并不被删除。

资源块的块标记为 5,记录对象的附属信息等。

由内容扰动模型可知,水印嵌入的最大幅度和海图内容可被改变的最大扰动区间 $m = \left\lfloor \frac{f}{2^b \times \Delta} \right\rfloor$; 由海图文件的组织结构可知,扰动区间主要位于海图的空间特征文件,进一步地,对象的坐标定义块内。

2.2 嵌入算法

通过对于扰动空间的定义,可以灵活地将水印信息嵌入到海图文件的数据块中,在保持稳健性能的同时获得较好的隐藏效果。具体算法如下所示:

输入:二维电子海图 V_0 , 二值水印图像 $W(n \times n)$; 输出:含水印的电子海图 V_w 。

```

(1)  $m = \left\lfloor \frac{f}{2^b \times \Delta} \right\rfloor, S = V_0$  拟嵌入水印的对象个数,  $T = \left\lfloor \frac{n \times n}{S} \right\rfloor$ 
(2) if  $T > S$  then 提示“水印过大,无法嵌入地图!”
(3) for  $i = 0$  to  $T - 1$  do
(4)      $W'_i = W_i \pm \frac{m - i \times \frac{m}{S}}{S}$ 
(5) end for
(6)  $L = V_0 =$  中定义块  $B[]$  总长度
(7) for  $j = 0$  to  $W'[]$  do
(8)     计算  $B[j]$  信息位长度  $l, W'[]$  长度  $a$ 
(9)     if  $a = 0$  then exit
(10)    else if  $L - a > l - 1$  then
(11)         $B[j][L - 1 - a \cdots L - 1] = W'[j]$ 
(12)    else  $j++$ 
(13)    end for
    
```

2.3 提取算法

对于水印信息的提取,则是对地图文件关键数据块的一次快速遍历,根据快标记识别和存储水印比特值。方法如下所示:

输入:含水印的二维电子海图 V' ; 输出:水印 W_i 。

```

(1)  $B[] =$  块标记为 2 的对象定义块,块大小为 1024 字节,  $S = V'$  被嵌入水印的对象个数
(2) for  $i = 0$  to  $S - 1$  do
(3)     if  $\text{int}(B[i][1023]) \geq 0$  then
(4)         { 将  $B[i][1023]$  按递增插入序列  $N[k_i], k_i$  是排序后插入值所处位置
(5)             for  $j > 0$  do
(6)                 if  $B[i][1023 - j]$  it not null & 不是定义结束符 then  $j++$ 
(7)                 else  $B[i][1022 \cdots 1023 - j]$  插入序列  $V_w[k_i]$ , 根据  $k_i$  调整序列  $V_w[]$ 
(8)             end for
(9)             else  $i++$ 
(10)            end for }
(11) for  $i = 0$  to  $\text{Len}(V_w[]) - 1$  do
(12)     计算  $W_i = W_i \mp \frac{m - i \times \frac{m}{S}}{S}$ 
(13) 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

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

其中 N 为水印长度。

3 实验与分析

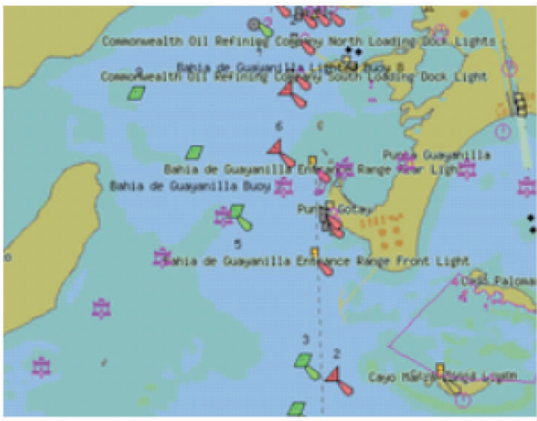
本文选择基于 MO 控件的 ArcGIS 环境。如图 3 所示,为实验用电子海图。

误码率:对于提取到的水印信息编码 W_i , 若 W_i 发生错误的概率为 $p(0 < p < \frac{1}{2})$, 则 W_i 与原始水

印 W_0 完全一致的概率为 $1 - p$ 。那么, W_i 与 W_0 存在编码差异的概率是 p , p 称为误码率。



(a) A



(b) B

图 3 测试用电子海图

对于水印标识采用统一的水印标记信息,如图 4 所示。



图 4 测试用数字水印

3.1 鲁棒性测试

为验证水印方案的鲁棒性,我们在对电子海图进行空间简化和数据压缩测试条件下进行检测,表 3 和表 4 分别给出了攻击后的误码率。同时,本文选取了另一种空间特征算法进行对比测试。

表 3 算法对地图简化攻击的鲁棒性

简化阈值/m	误 码 率	
	本文算法	文献[4]
0.2	0.012	0.02
0.4	0.031	0.03
0.6	0.044	0.05
0.8	0.067	0.06
1.0	0.082	0.09

表 4 算法对压缩攻击的鲁棒性

压缩算法	误 码 率	
	本文算法	文献[4]
文献[5],压缩比例 5%	0.022	0.02
文献[6],压缩比例 10%	0.031	0.07
文献[5],压缩比例 15%	0.039	0.05
文献[6],压缩比例 20%	0.068	0.08

本文采用的简化算法为道格拉斯—普克法,简化阈值在 0.2 ~ 1.0m,由表 3 可知,本文算法与文献[4]算法具有相当的抗简化攻击能力。

本文采用文献[5,6]的数据压缩算法,海图最高压缩比例达到 20%。随着压缩比例的提高,可以看到本文的误码率优于文献[4]算法,这主要是由于本文水印嵌入方法考虑到最大扰动区间和合理嵌入顺序,消除了随机性和嵌入不当造成的水印去除问题。

3.2 常规编辑操作测试

通过前述实验可知本方案对海图简化、压缩攻击都有较好的鲁棒性,而面对常规的海图编辑操作^[7-9],本文方案依然有效,表 5 分别给出了多种操

作情况下水印检测的误码率。

由于电子海图本身具有无损缩放、平移和旋转的特性,在对象的相对空间特征不变情况下,误码率为0。对于格式转换攻击,是指海图在不同应用软件和背景应用下,所进行的数据存储模式的改变,会造成数据精度、数据长度、关联关系的改变,会对数据造成一定的差异。由于格式转换属于一种数据扰动,故纳入内容扰动模型之内考虑。此外,基于文件组织结构以及空间特征的水印嵌入对象集合的选区也规避了格式转换后水印信息丢失的部分风险。

表 5 算法对常规操作的鲁棒性

攻击类型	误 码 率	
	本文算法	文献[4]
平移操作	0	0
旋转操作	0	0
缩放操作	0	0
格式转换	0.06	0.02
剪切 1/8	0.09	0.05
剪切 1/4	0.14	0.11
剪切 1/2	0.25	0.18

对于剪切攻击是多种水印算法都无法鲁棒的,主要是由于剪切操作的随机性,以及水印信息依赖于海图内容存在的本质特性,但本文相对文献[4]仍然具有较好的鲁棒性。

3.3 水印容量测试

为了测试本文引入的内容扰动模型对内容扰动幅度的判断能力,本文在人的视觉系统(HVS)无法检测水印嵌入的前提下,对水印容量进行了测试。

由于文献[4]的零水印方案缺少可比性,故选取了文献[10]和[11]所述的水印方法,分别是基于空间特征和属性特征的可逆水印算法。

如表6所示,HVS检测^[12]分三个阶段,即盲检测、提供原始海图文件参考检测、告知水印嵌入方法后的视觉检测。从图中可以看出,本文算法引入内容扰动模型后,水印嵌入对海图内容改变极其轻微,

在保持理想水印容量前提下,获得一个均衡的鲁棒性能。

表 6 水印容量测试

HVS 强度	水 印 容 量		
	本文算法	文献[10]	文献[11]
盲检测	20847	10984	17932
原始数据参照检测	20182	5920	7429
嵌入方法告知检测	18958	2092	3901

4 结 论

本文提出了一种基于二维电子海图内容扰动模型的可逆水印算法,通过内容特征模型构建数据扰动区间,并结合海图文件组织结构以及空间特征,构建了水印方案。实验验证了该方法具有较好的鲁棒性,并且整个过程中保持较好的水印容量,因此适用于公开传播环境下的二维电子海图的版权保护。

参考文献

[1] Mena J B. Antomatic vectorization of segmented road networks by geometrical and topological analysis of high solution binary images. *Knowledge-Based Systems*, 2006, 19 (1):704-718

[2] Voigt M, Gerrit S. A high capacity watermarking system for digital maps. In: *Proceedings of the 2004 Multimedia and Security Workshop on Multimedia and Security*, Magdeburg, Germany, 2004:180-186

[3] Peng F, Liu Y, Long M. Reversible watermarking for 2D CAD engineering graphics based on improved histogram shifting. *CAD Computer Aided Design*, 2014, 49:42-50

[4] 曹刘娟, 门朝光, 孙建国. 二维矢量地图双重零水印算法. *哈尔滨工程大学学报*, 2011, 32(3):340-344

[5] 李青元, 刘晓东, 曹代勇. Web GIS 矢量空间数据压缩方法探讨. *中国图形图像学报*, 2001, 6(12):1225-1229

[6] 李琦, 杨超伟, 陈爱军. WebGIS 中的地理关系数据库

模型研究. 中国图象图形学报,2000,5(2):119-123

[7] Bi N, Sun Q Y, Huang D R, et al. Robust image watermarking based on multiband wavelets and empirical mode decomposition. *IEEE Transaction on Image Process*, 2009, 16(8):195-1966

[8] Davydov A, Kovalev A, Lzyurov K. Distortion measure of watermarking 2D vector maps in the mesh-spectral domain. In: Proceedings of the 17th DSP 2011 International Conference on Digital Signal Processing, Corfu, Greece, 2011. 1123-1129

[9] Kim Junyeop. Vector map digital watermarking using an- gles. 6th International Conference on Networked Computing and Advanced Information Management, Jeju Island, Korea, 2010. 417-423

[10] 曹刘娟, 门朝光, 孙建国. 基于空间特征的二维矢量地图可逆水印算法原理, 测绘学报, 2010, 39(4):422-427

[11] 孙建国, 门朝光. 基于二维矢量地图属性特征的数字水印算法. 高技术通讯, 2009, 19(7):713-717

[12] Harrie L, Sarjakoski J. Simultaneous graphic generalization of vector datasets. *GeoInformatica, Computers Environment and Urbansystems*, 2012, 6(3):233-261

**Research on a reversible watermarking technique
for two dimension electronic charts
based on content perturbation**

Lai Mingzhu * **, Sun Jianguo *, Zhang Guoyin *

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

(** Department of Software, Harbin University of Science and Technology, Harbin 150080)

Abstract

A technique for reversible watermarking two-dimensional electronic charts was studied to solve the problem that conventional watermark embedding strategies can not achieve the practical watermarking for two-dimensional electronic charts because they often brings chart content disturbances, so it is difficult to meet the charts' requirements of high fidelity and high content integrity. On the one hand, a model for content perturbation control was obtained based on the research on the chart content features. On the other hand, a reversible watermarking scheme based on chart data's organizational structure and spatial characteristics was proposed to improve the practicality of electronic charts. The experiment and study show that this method is very robust. Compared with the similar methods, this method can adapt to the data perturbation brought by conventional chart operation, and has the better robustness.

Key words: electronic chart, information hiding, content perturbation, digital watermarking, reversibility