

矢量地图水印归一化相关检测的 可行性分析与改进

钟尚平^(1,2) 高庆狮⁽¹⁾

⁽¹⁾ (中国科学院计算技术研究所软件室, 北京 100080) ⁽²⁾ (中国科学院研究生院, 北京 100039)

摘 要 为了使矢量地图水印检测算法具有鲁棒性和运算简单, 首先分析了矢量地图归一化相关检测算法的可行性: 鉴于矢量地图图形的可嵌入水印长度较短, 而且其 DFT 系数的幅度不能较好地满足高斯分布等原因, 对不同图形的水印, 采用归一化相关检测方法无法确定检测阈值, 因此是不可行的, 同时还分析了图形简化对水印检测的影响。为使水印算法抗简化, 对几何失真鲁棒及减小运算量, 提出了以图形的特征点的 DFT 系数幅度和相位作为水印嵌入域的方法, 并提出了一种以提取水印与原始水印的相关系数作为检测值的实用的检测算法。实验结果表明, 该算法是鲁棒的和可行的。

关键词 矢量地图水印 离散傅里叶变换 归一化相关检测 可行性分析 可行鲁棒水印算法

中图法分类号: TP309 **文献标识码:** A **文章编号:** 1006-8961(2006)03-0401-09

The Feasibility Analysis of Normalized-correlation-based Vector Maps Watermarking Detection Algorithm and the Improved Watermarking Algorithm

ZHONG Shang-ping^(1,2), GAO Qing-shi⁽¹⁾

⁽¹⁾ (Software Division, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

⁽²⁾ (Graduate School of the Chinese Academy of Sciences, Beijing 100039)

Abstract Because the watermark lengths of many applied vector maps' polygonal lines are short, and the Gaussian distribution is not suited to describe magnitudes of discrete Fourier transform coefficients for vector maps' polygonal lines to difference polygonal lines in the vector maps, the uniform threshold of normalized-correlation-based watermarking detection is hardly confirmed. In this paper, to improve algorithm's efficiency and be invariant to simplification, the discrete Fourier transform is only used for the feature points of vector maps. The watermarking algorithm embeds the watermarks by changing both the magnitudes and the phases of discrete Fourier transform mid-frequency coefficients for vector maps' polygonal lines. The correlation coefficient of the detected watermarks and the embedded watermarks is regarded as the watermark detection value. Given applied map data, computing results show that watermarks generated by this technique are robust, and the algorithm is fairly good in efficiency and Feasibility.

Keywords vector map watermark, discrete Fourier transform (DFT), normalized correlation detection, feasibility analysis, feasible robust watermarking algorithm

1 引 言

矢量图形数据作为多媒体信息的一种重要表现

形式, 其相应的版权保护及数据认证等数字水印技术近年来逐渐得到了关注。但相对于人们对图像、视频及音频数据水印技术的广泛关注, 矢量图形数据, 特别是 2 维矢量图形数据的水印技术见诸于文

基金项目: 国家自然科学基金项目(60273016); 中国科学院计算技术研究所基金项目(20016250)

收稿日期: 2004-07-12; 改回日期: 2005-04-15

第一作者简介: 钟尚平(1969 ~), 男。2005 年获中国科学院计算技术研究所, 中国科学院科学研究生院博士学位, 现为福州大学讲师。主要研究方向为网络信息安全、算法设计等。E-mail: spzhong99@21cn.com; zhongshangping@software.ict.ac.cn

献的研究工作报导还很少。2 维矢量图形不仅在地理信息系统 (GIS)、艺术设计及动画制作等领域具有广泛的应用,而且随着信息技术的飞速发展,以“数字地球”为契机,数字地图的应用可谓方兴未艾。对大量的 2 维矢量地图数字水印技术的研究无疑是有较大意义的。

2 维矢量图形是一种在数据表示,使用方式等方面有着其自身特点的数据形式,其与以栅格点阵形式存在的 2 维图像和一般由三角面片网格组成的 3 维几何模型有较大的不同。人们很难简单地应用现有的数字水印技术来处理 2 维矢量图形数据^[1]。文献[1]列出了所有两篇有关 2 维图形数字水印技术的参考文献[2,3]。文献[3]提出了一个通过重新参数化的方法来对 NURBS (non-uniform rational B-spline) 曲线进行数据嵌入,适合于 3 维及 2 维图形数据的数字水印嵌入。文献[2]提出了一种基于 Fourier 描述子的针对封闭多边形数据的数字水印算法。该算法充分利用了几何变换下 Fourier 描述子的性质,使水印嵌入在矢量图形 DFT (discrete Fourier transform) 系数的幅度上,同时构造了对于几何变换不变的归一化检测统计量,并采用了线性相关检测方法。文献[4]以及文中提到的文献[5~7]是目前能查询到的除文献[2,3]外,针对矢量地图的水印算法的参考文献,文献[4]算法是文献[5]算法的改进,它是先把整块地图的坐标点连成 Delaunay 三角网格,并把三角网格转化成 Laplacian 频谱,然后将水印信息嵌入在 Laplacian 频谱系数中。文献[6]是把矢量地图转化成类似栅格图像表示方式,并采用了类似基于小波变换域的图像水印算法。文献[7]则是把水印信息嵌入到图形的个别点的坐标值上,但按文献[4]的评价,其作为空间域水印算法对随机噪声攻击的鲁棒性是很差的。总之,尽管针对 2 维矢量地图的数字水印技术的研究刚刚兴起,但大量的针对 3 维矢量图形和其他媒体的数字水印技术的研究结果是可借鉴的对象。2 维矢量地图的数字水印算法目前主要有基于几何数据变换域^[2,3,6]的方法、基于拓扑关系图的 Laplacian 频谱域^[4,5]的方法以及基于空间域的水印算法^[7]等。

考虑到实际的城市交通矢量地图(如 1:10 000 比例尺的厦门市交通图,MAPINFO 格式)大多图层中图形的可嵌入水印长度较短,特别当为了使矢量地图水印具备一定的抗简化性能及为了减小算法运算量,只对图形的特征点嵌入水印时,可嵌入水

印长度很短。依据信号检测原理^[8],若噪声环境偏离了原设计的最佳检测器的假定(假定噪声为高斯分布),则最佳线性检测器的性能将严重退化,可是据文献[9]以及本文的实验结果表明,高斯分布不适于描述矢量地图封闭曲线 DFT 系数的幅度。归一化相关检测方法是以可嵌入水印长度在统计意义上足够长和水印嵌入能域较好地满足高斯分布为假设条件而建立的,但对于实际的矢量地图,这些假设条件并不能很好地满足。实验结果表明,对于矢量地图中不同图形的水印,由于无论是相关检测,还是归一化相关检测都难以确定检测阈值,因此是不可行的。而且,具体计算时,也无法用同一阈值来生成不同长度的均值为零的二值水印序列,而水印序列的均值为零则是相关检测所要求的。本文在进行上述分析的基础上,考虑到图形简化对水印检测的影响,为使水印算法抗简化,并对几何失真鲁棒及减小运算量,提出了如下算法,即对矢量地图的特征点的 DFT 系数幅度和相位同时嵌入水印,由于二值水印序列的均值可不必为零,以致于不同长度水印可用同一阈值生成,因此本文采用了提取水印与原始水印的相关系数作为检测值的实用的检测方法。这虽不是一个盲检测方法,但作为对矢量地图的版权保护是可行的,且具有较好的鲁棒性。本文算法对其他类型矢量地图及矢量图形也有一定的指导意义或基本可以直接使用。

2 基于 DFT 的矢量地图水印相关检测的可行性分析

文献[2]中引用的 Fourier 描述子^[10]定义如下:

经采样的某一多边形曲线可表示为: $u(n) = x(n) + iy(n)$, $n = 0, 1, \dots, N-1$, i 为虚数单位。使用式(1)就可计算出这 N 个点的离散傅里叶变换系数 $a(k)$, 而 $a(k)$ 即作为 Fourier 描述子。

$$a(k) = \sum_{n=0}^{N-1} u(n) \exp\left(-\frac{i2\pi kn}{N}\right) \quad (1)$$

其中

$$u(n) = \frac{1}{N} \sum_{k=0}^{N-1} a(k) \exp\left(\frac{i2\pi kn}{N}\right) \quad (2)$$

地理信息系统中常见的地图图形数据结构为矢量结构,即通过记录坐标的方式来尽可能精确地表示点、线、多边形等地理实体。对于点实体,矢量结构中只记录其在特定坐标下的坐标和属性代码;对

于曲线实体,矢量结构中,只记录逼近曲线的足够多的小线段的端点坐标,并将曲线表示为一个坐标序列。“多边形”在地理信息系统中是指一个任意形状,边界完全闭合的空间区域。把这样的闭合区域称为多边形的原因是,由于区域的边界线同前面介绍的线实体一样,可以被看作是由一系列多而短的直线段组成,且每个小线段可作为这个区域的一条边,因此这种区域就可看作是由这些边组成的多边形了^[11]。事实上,基于文献[2]思想的图形水印算法就适合于矢量地图的多边形实体及曲线实体。由于类似城市交通地图的多边形实体与曲线实体相比可允许嵌入更大的水印能量,因此本文仅讨论矢量地图多边形实体(封闭曲线)的水印算法。

设由封闭曲线 DFT 中频系数幅度或相位构成的盲水印宿主序列为 $\{x(k)\}$, 水印序列为 $\{w(k) | w(k) \text{ 为 } 1 \text{ 或 } -1\}$ (可由伪随机数发生器或混沌映射生成), 待测序列为 $\{y(k)\}$, $k=0, 1, \dots, N-1$ 。另外,还可将盲水印检测问题转换为如下假设检验问题:

$$H_0: y(k) = x(k), k=0, 1, \dots, N-1 \text{ (无水印)} \quad (3)$$

$$H_1: y(k) = x(k) + \lambda w(k)x(k), k=0, 1, \dots, N-1 \text{ (有水印, } 0 < \lambda \leq 1 \text{ 是控制水印强度的参数)} \quad (4)$$

设 $c = \sum_{k=0}^{N-1} w(k)y(k)$ 为水印序列与待测序列的相关值(以下简称相关值)。按水印检测原理,通常通过将相关值与预定阈值相比较来检测水印的存在与否,而且在加性高斯噪声的作用下,它是信号(水印)检测的最优方法。在水印盲检测算法中,变换域作为噪声,而水印系列作为信号。对于不同的待测水印序列,由于相关值 c 的大小差异较大,所以,无法确定统一的预定阈值。为此,文献[2]构造了下述归一化相关值统计量 c_{normal} 。设 $w(k)$ ($k=0, 1, \dots, N-1$) 为独立同分布的随机变量($w(k)$ 的均值为零), $x(k)$ 也为独立同分布的随机变量,且 $w(k)$ 与 $x(k)$ 是独立的,则 $c = \sum_{k=0}^{N-1} w(k)y(k)$ 的均值 $E(c)$ 为:当 $w(k)$ 为正确水印时, $E(c) = N\lambda E(x(k))$;而当 $w(k)$ 为非正确水印或无水印时, $E(c) = 0$ 。又因 $E(y(k)) = E(x(k) + \lambda w(k)x(k)) = E(x(k)) + \lambda E(w(k))E(x(k)) = E(x(k))$, 故当 $w(k)$ 为正确水印时,则 $E(c) = N\lambda E(y(k))$ 。若构造的归一化相关值统计量为 $c_{\text{normal}} = \frac{c}{N\lambda E(y(k))}$, 则当 $w(k)$ 为正确

水印时, $c_{\text{normal}} > T$, 此时检测门限 T 趋于 1; 而当 $w(k)$ 为非正确水印或无水印时, 相关值 c 趋于 0。当 DFT 中频系数幅度为水印嵌入域时, 则此归一化相关值统计量对平移、旋转及缩放等几何失真免疫的^[2]。

$$\text{由 } c = \sum_{k=0}^{N-1} w(k)y(k) = \sum_{k=0}^{N-1} w(k)x(k) + \lambda \sum_{k=0}^{N-1} w^2(k)x(k)$$

可见, 只有当 $\lambda \sum_{k=0}^{N-1} w^2(k)x(k)$ 相对较大, 而 $\sum_{k=0}^{N-1} w(k)x(k)$ 很小(其中 $w(k)$ 为正确水印), 以至于可忽略不计时, $E(c) = N\lambda E(x(k))$ 才成立, 进而可用上述趋于 1 的检测门限 T 来检测有无水印; 而要使 $\lambda \sum_{k=0}^{N-1} w^2(k)x(k)$ 相对较大, 同时使 $\sum_{k=0}^{N-1} w(k)x(k)$ 很小, 以至于可忽略不计的条件为: (1) 可嵌入水印长度 N 在统计意义上足够大; (2) $x(k)$ 较好地满足高斯分布; (3) $w(k)$ 的均值为零; (4) λ 不能太小。下面对这 4 个条件进行分析。

2.1 矢量地图封闭曲线的可嵌入水印长度

采用实际使用的厦门市城市交通矢量地图嵌入水印时, 首先对地图的所有 4 种多边形类图层: 水系(面)、区域、绿地范围、居民地(面)图形的坐标点数进行统计。统计结果表明: 区域图层中只有少数图形的坐标点数有上 1 000 个点, 水系(面)和绿地范围图层中的图形均为 1 000 个点以内, 多数为 200 个坐标点数以内, 而居民地(面)图层中图形都在 100 个坐标点数以内。因此, 这种矢量地图不同于图像数据(如: 256×256 或更大)具有统计意义上较大的可嵌入水印长度, 实际使用的矢量地图, 其可嵌入水印的长度在统计意义上较小。特别当为了使矢量地图水印具备一定的抗简化性能和为了减小算法运算量, 并使水印具备透明性和鲁棒性, 只对图形的特征点的 DFT 中频系数幅度或相位嵌入水印时, 由于可嵌入水印长度很短, 因此, 矢量地图可嵌入水印长度难以满足相关检测所要求的在统计意义上足够长的条件。关于矢量地图曲线简化方法及水印算法的抗简化性能分析将在 3.1 节介绍。

2.2 矢量地图封闭曲线 DFT 系数幅度与相位的高斯分布“K-S”检验结果

与基于变换域的图像水印嵌入算法^[12]一样, 为了使水印具备透明性和鲁棒性, 基于 DFT 的封闭曲线水印算法是把水印系列嵌入在矢量图形 DFT 中频系数的幅度或相位上。下面计算矢量地图封闭曲线 DFT 中频系数幅度和相位的高斯分布“K-S”检

验结果值。

“K-S”检验^[13]是利用 Kolmogorov 定理中的统计量： $\sqrt{n}D_n = \sqrt{n} \sup_{-\infty < x < +\infty} |F_n(\hat{x}) - F(\hat{x})|$ 的极限分布 $K(\hat{x})$ (上式中 $F_n(\hat{x})$ 为随机变量 $\hat{X}$ (或总体) 的经验分布函数) 来对假设问题 $H_0: \hat{X}$ 具有分布函数 $F(\hat{x}) = F_0(\hat{x})$ 进行显著性检验。 H_0 的拒绝条件为:

$$D_n > D_{n,\alpha} = \frac{k_\alpha(n)}{\sqrt{n}}, \text{ 其中, } n \text{ 为样本观测次数, } \alpha \text{ 为显著性水平, } k_\alpha(n) \text{ 可通过 Kolmogorov 分布表查得。}$$

特别当 n 较大时, 可近似地决定 $D_{n,\alpha}$ 的值, 即有

$$D_{n,0.01} \approx 1.628/\sqrt{n}, D_{n,0.05} \approx 1.358/\sqrt{n}。 \text{ 在具体计算 } D_n \text{ 时, 由于 } F_n(\hat{x}) \text{ 为阶梯函数, 故 } D_n = \sup_{-\infty < x < +\infty} |F_n(\hat{x}) - F(\hat{x})| = \max_{1 \leq k \leq n} |F_n(\hat{x}_k) - F_0(\hat{x}_k)|,$$

$|F_n(\hat{x}_{k+1}) - F_0(\hat{x}_k)|$, 其中 $F_n(\hat{x}_{n+1})$ 可理解为 1, 而 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$ 则为次序统计量 $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n$ 的观测值。若 $D_n > D_{n,\alpha}$ 则拒绝 H_0 ; 否则接受 H_0 , 即认为 X 具有分布函数 $F(\hat{x}) = F_0(\hat{x})$ 。

将上述“K-S”检验方法应用于厦门市城市交通地图, 以便检验 DFT 中频系数相位与幅度能否较好地满足高斯分布。本文选取了实际地图多边形类图层的 3 种图形(水系(面)、区域、绿地范围), 按图形复杂层次随机抽取各图层中的 3 个图形作为实验数据。另外, 由于实际地图中的“居民地(面)”图形较简单(坐标点数很少), 因此不作为实验数据。

矢量地图封闭曲线 DFT 中频系数的相位与幅度高斯分布的“K-S”检验结果见表 1。

表 1 矢量地图封闭曲线 DFT 中频系数的相位与幅度“K-S”的检验结果

Tab.1 “K-S” detection values of magnitudes and phases of discrete Fourier transform mid-frequency coefficients for vector maps' polygonal lines

图层	图形类别及编号	坐标点数	取中频数	$D_{n,0.01}$ 值	$D_{n,0.05}$ 值	相位 D_n 值	幅度 D_n 值	
多边形图层	水系(面)	1	407	150	0.132 926	0.110 880	0.072 226	0.226 614
		2	179	70	0.194 583	0.162 312	0.114 392	0.256 282
		3	83	40	0.257 409	0.214 719	0.088 379	0.264 595
	区域	1	1049	300	0.093 993	0.078 404	0.061 477	0.190 040
		2	131	50	0.230 234	0.192 050	0.079 891	0.226 609
		3	74	40	0.257 409	0.214 719	0.099 309	0.209 619
	绿地范围	1	216	80	0.182 016	0.151 829	0.121 643	0.163 145
		2	161	60	0.210 174	0.175 317	0.135 925	0.185 964
		3	72	40	0.257 409	0.214 719	0.106 764	0.240 839

由表 1 可知, 矢量地图封闭曲线 DFT 中频系数的幅度不能很好地满足高斯分布, 相对而言, 相位比幅度能更好地满足高斯分布。对简化的封闭曲线也有此结果。注: 以上相位值在 $[0, 2\pi]$ 间取值, 以下同。

2.3 关于均值为零的二值水印序列

实际计算时, 均值为零的二值水印序列 $\{w(k)\}$ 一般的生成方法为: 先由伪随机数发生器或混沌映射生成“0~1”间均匀分布的随机实数, 再通过某一给定阈值(如: 0.5 左右的值), 把大于阈值的随机实数定为 1, 小于阈值为 -1, 或相反, 从而生成均值为零的 2 值水印序列 $\{w(k)\}$ 。另外, 也可以用一组阈值, 如: 用 10 个阈值把 0~1 近似地 10 等分, 同一区间的随机实数为相同水印值, 相邻区间的随机实数定为不同

水印值。显然, 不同长度的均值为零的二值水印序列一般不能由同一阈值(或一组阈值)来生成。

2.4 λ 的大小

在水印透明性范围内, 对于矢量地图幅度水印, 嵌入强度参数 λ 越大, 水印相关检测值越大; 但对于相位水印, λ 增大到一定程度, 就有可能使嵌入水印的 DFT 系数跨越原来所处的象限, 使水印相关检测值的大小难以预料。

综上所述: (1) 对实际使用的城市交通矢量地图中不同图形水印使用统一的归一化相关检测阈值的条件不能很好地满足; (2) 由于对不同长度的均值为零的二值水印序列, 一般不能由同一阈值(或一组阈值)来生成; 对矢量地图不同图形中嵌入的水印, 需要确定不同的阈值, 增加了水印算法的复杂

性。这两个问题,影响了基于 DFT 的矢量地图水印算法的实用性和简单性。就表 2 中的 6 个图形,用同一初值生成不同长度的均值为零的二值水印,在不作变换及不被任何恶意攻击的环境下,对各图形的幅度水印(水印嵌入强度参数 $\lambda = 0.1$)和相位水印($\lambda = 0.015 \sim 0.09$)进行相关盲检测,得到如表 3 所示的结果(为便于比较,无水印时, c_{normal} 的检测值为 c 除以 $N\lambda E(y(k))$ 后的值)。

表 2 矢量地图图形的幅度和相位水印相关盲检测结果

Tab.2 Blind watermarking detection values of magnitudes and phases for vector maps' polygonal lines

图层	图形类别及编号	坐标 点数	取中 频数	幅度 c 值		幅度 c_{normal} 值		幅度 c 值		幅度 c_{normal} 值		
				有水印	无水印	有水印	无水印	有水印	无水印	有水印	无水印	
多边形图层	水系(面)	1	407	150	0.002647	-0.016763	0.141271	-0.863609	19.05526	12.04489	2.717103	1.718154
		2	179	70	0.003672	-0.002762	0.580647	-0.429324	-7.202172	-14.16389	-1.036438	-2.034540
	区域	1	1049	300	1.021519	0.285303	1.366347	0.387527	62.48366	33.02369	2.118831	1.120969
		2	131	50	0.002190	0.000416	1.223123	0.234601	11.38777	-3.577035	0.741653	-0.232962
	绿地范围	1	216	80	0.013821	-0.002288	0.862863	-0.142040	9.470677	2.236967	1.308878	0.309242
		2	161	60	0.003680	-0.001256	0.753183	-0.254484	19.03244	9.162479	1.902488	0.915883

表 2 的结果表明, $\sum_{k=0}^{N-1} w(k)x(k)$ 相对于 $\lambda \sum_{k=0}^{N-1} w^2(k)x(k)$ 不能忽略不计,事实上,“有水印 c_{normal} 值”减去“无水印 c_{normal} 值”约等于 1。由于用统计量 c 及 c_{normal} 来检测水印,对不同图形无法确定检测阈值,因此难以实用。

3 可行并鲁棒的矢量地图水印算法

本文考虑到图形简化对水印检测的影响,为使水印算法抗简化,对几何失真鲁棒及减小运算量,提出了如下水印嵌入算法,对矢量地图的特征点的 DFT 系数幅度和相位同时嵌入水印。并且,采用了提取水印与原始水印的相关系数作为检测值的实用的水印检测方法,这样,二值水印序列的均值可不必为零,以至于不同长度水印可用同一阈值生成。图 1 为本文矢量地图水印算法流程图。

3.1 曲线简化方法介绍与曲线简化对水印的影响分析

矢量地图的存储特点决定了大比例尺矢量地图经过曲线简化处理,依然可进行广泛应用,如: WebGIS 应用(目前数字地图的大数据量依然是制约 WebGIS 应用的主要因素)。另外,大比例尺矢量地图可通过简化、移位、缩放等地图综合操作生成小比例尺矢量地图^[14]。另一方面,若只对曲线的特征点作离散傅里叶变换,可以提高水印算法的计算

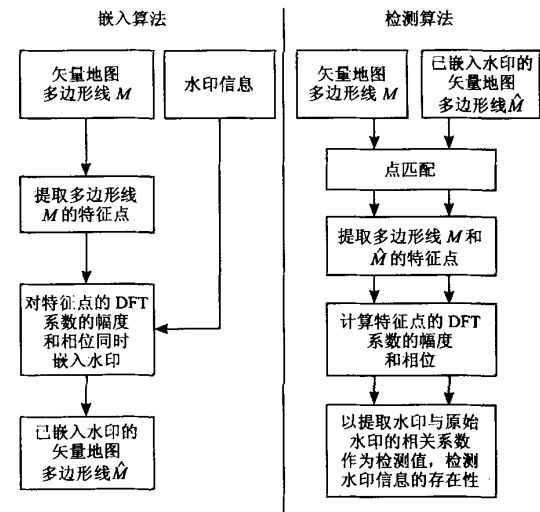


图 1 本文矢量地图水印算法流程图

Fig.1 An overview of this paper's watermark embedding and extraction scheme

效率(计算一个长度为 N 的 1 维 DFT,需要作 N^2 次复数乘法和 $N(N-1)$ 次复数加法运算,当 N 很大时,由于运算量巨大,因此有了复数乘法和复数加法正比于 $N\ln N$ 的 FFT(fast Fourier transform)算法。由此可见,为了使矢量地图水印算法具有更好的版权保护功能和较小的算法运算量,有必要介绍曲线简化方法,并就曲线简化对水印的影响进行分析。

曲线的简化算法有多种,本文采用文献[15]中

的算法,其原理是提取曲线上局部区域的最大曲率点作为近似多边形的特征点,以特征点构成简化曲线。该算法的具体步骤如下:

(1) //初始化,从 p_0 点开始提取所有的特征点
let $i=0$; let $j=i+2$; //let p_0 点作为第 1 个特征点;

(2) //在 (p_i, p_j) 区间的曲线上,寻找到直线 $p_i p_j$ 距离最大的点 p_m

If $j > N$ then stop, else find a point p_m , 满足
 $d(p_i p_j, p_m) = \max_{i < k < j} d(p_i p_j, p_k)$;

(3) //判断 p_m 是否为特征点

If $d(p_i p_j, p_k) > D$ then 提取 p_m 作为新的特征点; Let $i=m$; let $j=i+2$; 转步骤 2, else $j=i+1$; 转步骤 2。

这里,

$$d(p_i p_j, p_m) = \left| \frac{(y_j - y_i)x_m - (x_j - x_i)y_m + (y_i x_j - x_i y_j)}{((x_j - x_i)^2 + (y_j - y_i)^2)^{1/2}} \right|$$

表示点 p_m 到直线 $p_i p_j$ 的距离。由于简化图形是原图形的很好的近似,因此可以通过改变上算法中的参数 D 来控制图形的简化程度。

现假设对封闭曲线的第 $p, p+1, \dots, q$ 个 DFT 系数(幅度或相位)嵌入水印,又设 $W_N^{-kn} = e^{-2\pi kni/N}$, i 为虚数单位,则由式(1),(2)可得到以下带水印曲线:

$$\hat{u}(n) = \frac{1}{N} \times \sum_{k=0}^{p-1} a(k) W_N^{kn} + \frac{1}{N} \sum_{k=p}^q \hat{a}(k) W_N^{kn} + \frac{1}{N} \sum_{k=q+1}^{N-1} a(k) W_N^{kn}$$

这里不妨设

$$\hat{a}(k) = a(k) + \lambda w(k) a(k)$$

下面对带水印曲线未被简化处理和被简化处理时的水印情形进行分析:

(1) 若带水印曲线未被简化处理,即保持原有的 N 个带水印点,则有:当 $k_1 = p, p+1, \dots, q$ 时,

$$\begin{aligned} \hat{a}(k_1) &= \sum_{n=0}^{N-1} \hat{u}(n) W_N^{-k_1 n} = \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{k=0}^{p-1} a(k) W_N^{kn} \right) W_N^{-k_1 n} + \\ &\quad \frac{1}{N} \lambda \sum_{n=0}^{N-1} \left(\sum_{k=p}^q w(k) a(k) W_N^{kn} \right) W_N^{-k_1 n} \\ &= a(k_1) + \lambda w(k_1) a(k_1) \end{aligned}$$

这个等式应用定理 1^[16] 不难推出。

定理 1 如果 k 为不等于 0 或 N 的任意整数, 则 $\sum_{n=0}^{N-1} W_N^{kn} = 0$ 。其证明见文献[16]。

综上所述,若带水印曲线未被简化处理,则水印

检测时,由 $\hat{a}(k_1) = a(k_1) + \lambda w(k_1) a(k_1)$ 即可完全区分出(不带噪声的理想情形下)水印项 $\lambda w(k_1) a(k_1)$,进而即可完全区分出幅度或相位水印。

(2) 若带水印曲线已被简化处理,为了表述简单,不妨设简化掉一个点 $\hat{u}(N-1)$, 则有

$$\begin{aligned} \hat{a}(k_1) &= \sum_{n=0}^{N-2} \hat{u}(n) W_N^{-k_1 n} = \frac{1}{N} \sum_{n=0}^{N-2} \left(\sum_{k=0}^{N-1} a(k) W_N^{kn} \right) W_N^{-k_1 n} + \\ &\quad \frac{1}{N} \lambda \sum_{n=0}^{N-2} \left(\sum_{k=p}^q w(k) a(k) W_N^{kn} \right) W_N^{-k_1 n} \\ &= \frac{1}{N} \sum_{n=0}^{N-2} \left(\sum_{k=0}^{N-1} a(k) W_N^{kn} \right) W_N^{-k_1 n} + \\ &\quad \frac{1}{N} \lambda \sum_{n=0}^{N-2} \left(w(p) a(p) \frac{W_N^{n(p-k_1)}}{W_N^{nq/N(N-1)}} + \dots + \right. \\ &\quad \left. w(q) a(q) \frac{W_N^{n(q-k_1)}}{W_N^{nq/N(N-1)}} \right) \end{aligned}$$

由于 $\sum_{n=0}^{N-2} \frac{W_N^{n(k-k_1)}}{W_N^{nq/N(N-1)}}$ 对 $k=p, \dots, k_1-1, k_1+1, \dots, q$

一般不为零,所以 $\hat{a}(k_1) = a(k_1) + \lambda w(k_1) a(k_1)$ 一般不成立,由于其已难以区分出水印,从而难以检测出水印,第 4 章的实验结果也说明了这点。

易见,若先提取矢量地图封闭曲线的特征点,再对特征点做离散傅里叶变换,并以此变换的系数的幅度或相位作为嵌入域,则其就具有一定的抗简化性,因为按矢量地图的简化原则,必须保持图形的特征点,否则,将给图形带来较大失真,以至于使图形不具应用价值。

3.2 水印的嵌入与检测

水印按式(4)方式嵌入。对矢量地图图形的特征点的 DFT 系数幅度和相位同时嵌入水印。

如第 2 章所述,设水印宿主序列为 $\{x(k)\}$, 原始水印序列为 $\{w(k)\}$ 及待测序列为 $\{y(k)\}$, 则按下式提取的待测序列的水印系数为 $\hat{w}(k) = \text{sgn}(y(k) - x(k))$, sgn 为符号函数。当然,水印经过某些变换或人为的恶意攻击后,由于提取出的水印系数 $\hat{w}(k)$ 可能不会与嵌入的原始水印系数 $w(k)$ 完全相同,因此本文采用提取的水印 $\hat{w}$ 的值 $\hat{w}(k)$ 与原始水印 w 的值 $w(k)$ 的相关系数作为水印检测值来进行检测,其公式如下:

$$\begin{aligned} \text{Cov}(w, \hat{w}) &= \frac{\sum_{k=0}^{N-1} (w(k) - E(w))(\hat{w}(k) - E(\hat{w}))}{\sqrt{\sum_{k=0}^{N-1} (w(k) - E(w))^2} \sqrt{\sum_{k=0}^{N-1} (\hat{w}(k) - E(\hat{w}))^2}} \\ \text{Cov}(w, \hat{w}) &\text{取值在 } [-1, 1] \text{ 之间。如果 } \text{Cov}(w, \end{aligned}$$

$\hat{w}$ 超过某一检测阈值 T , 则判定此矢量地图图形中存在水印(在不作旋转和缩小变换及不被任何恶意攻击的环境下, 正确水印的 $Cov(w, \hat{w})$ 值为 1(或接近 1), 无水印的 $Cov(w, \hat{w})$ 值为 0(或接近 0))。检测阈值 T 的确定方法见第 4 章, 其与虚警概率的关系可参考文献[2, 17]。

采用提取水印 $\hat{w}(k)$ 与原始水印 $w(k)$ 的相关系数作为水印判决标准, 为一个非盲检测方法, 但该方法有以 c 和 c_{normal} 作为检测统计量的相关检测方法不具有的实用性, 即: (1) 对不同的矢量地图图形可用趋于 1 的检测阈值来检测; (2) 水印嵌入长度对水印检测影响不大; (3) 不用考虑水印嵌入域是否较好地满足高斯分布; (4) 不要求水印序列的均值为零, 因此, 可通过给定的初值和阈值(如: 0.5)来生成不同长度的二值水印, 即嵌入矢量地图中的不同图形; (5) 在保证水印透明性的条件下, 对幅度和相位水印, 嵌入强度参数 λ 可取固定值。

3.3 水印算法的鲁棒性分析

(1) 简化处理: 只对矢量地图的特征点嵌入水印, 而对矢量地图的非特征点不嵌入水印, 即使非特征点被简化掉, 将不影响水印检测, 从而具有一定的抗简化性。

(2) 平移变换: 由表 1 可知, 平移只影响第 1 个 DFT 系数, 本文以 DFT 中频系数作为嵌入域, 故平移对幅度和相位水印均无影响。

(3) 缩小、放大: 此变换仅改变 DFT 系数的幅度, 不改变相位, 故缩放对相位水印无影响。

(4) 旋转变换, 改变起始点及反射变换: 一般而言, 对矢量地图作旋转及反射变换后, 实际意义不大。其实, 由表 1 可见, 这些变换仅改变 DFT 系数的相位, 而并不改变幅度, 故这些变换对幅度水印无影响。

因此, 对矢量地图图形的特征点的 DFT 系数幅度和相位同时嵌入水印, 具有较好的鲁棒性。第 4 章实验结果也表明了这点。

4 实验结果

本文选取了表 1 中坐标点数为 1 049 的区域图形, 提取 367 个特征点(参数 $D=0.000\ 001\ 8$), 通过对 200 个 DFT 中频系数幅度和相位嵌入水印来对第 3 章两个改进的水印检测算法进行仿真实验。图 2 为 367 个特征点构成的图形; 图 3 为嵌入水印

后的图形(幅度水印 $\lambda=0.05$, 相位水印 $\lambda=0.02$, λ 是控制水印强度的参数)。矢量地图嵌入水印后基本是透明的。另外, 这两图是为了减小作图的工作量, 只由 367 个特征点构成的图, 实际图形由 1 049 个点构成将有更好的光滑性。算法对其他地图图形也有类似结果, 限于篇幅, 不能一一列举。

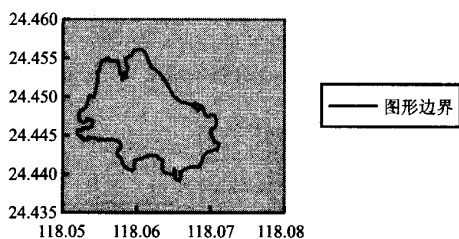


图 2 367 个特征点构成的图形

Fig. 2 A vector map polygonal line made up of 367 feature points

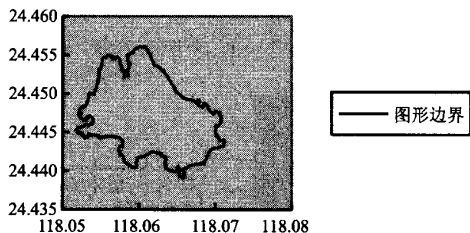


图 3 嵌入水印后的图形

Fig. 3 The watermarked polygonal line

4.1 检测阈值的确定

随机生成 10 000 个长度为 200 的水印序列, 其与原始水印的相关系数值在 $[-0.339\ 266, 0.335\ 151]$ 之间, 而正确的水印与原始水印的相关系数值为 1(或接近 1)。因此, 可以将检测阈值 T 定为 0.4, 也可以将其放大, 以降低系统发生虚检的概率。依此方法可确定其他长度水印序列的检测阈值。

4.2 简化对水印检测的影响

对已嵌入水印的图形假设简化掉一个点(如第 3 点)有如图 4, 图 5 检测结果。证实了 3.1 节的结论。

4.3 几何失真变换下的水印检测

图 6 为幅度水印不作变换与旋转变换(旋转 10°)后的检测图。图 7 为相位水印不作变换与缩小变换(表 1 中 $\alpha=0.8$)后的检测图。图 6、图 7 所示的实验结果表明, 旋转变换对幅度水印是免疫的, 而缩小变换则对相位水印是免疫的。水印对其他几何失真变换鲁棒性的实验结果与 3.3 节的分析完全一

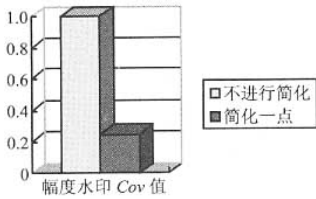


图 4 简化对幅度水印检测的影响

Fig. 4 The two magnitude-watermark detection values before and after the watermarked polygonal line is simplified

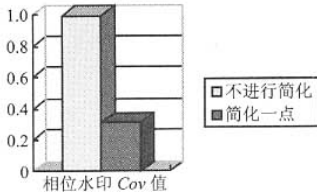


图 5 简化对相位水印检测的影响

Fig. 5 The two phase-watermark detection values before and after the watermarked polygonal line is simplified

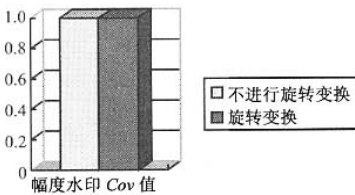


图 6 幅度水印不进行缩小旋转变换与旋转变换后的检测图

Fig. 6 The two magnitude-watermark detection values before and after the watermarked polygonal line is rotated

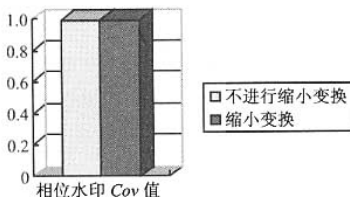


图 7 相位水印不进行缩小变换与缩小变换后的检测图

Fig. 7 The two phase-watermark detection values before and after the watermarked polygonal line is down-scaled

致, 限于篇幅, 不能一一列举。

5 结 论

本文首先从如下几方面分析了相关检测和归一化相关检测算法的可行性: (1) 指出了对于不同的

待测序列, 相关检测值的大小差异较大, 所以, 无法确定统一的预定阈值。(2) 提出了归一化相关检测成立的条件; (3) 指出了矢量地图图形的实际可嵌入水印长度较短, 难以满足归一化相关检测所要求的在统计意义上足够长的条件; (4) 用“K-S”检验方法得出了矢量地图封闭曲线 DFT 系数幅度不能较好地满足高斯分布的结果; (5) 具体计算时, 无法用同一阈值生成不同长度的均值为零的二值水印序列, 而水印序列的均值为零却是相关检测所要求的; (6) 实验结果表明了采用相关检测和归一化相关检测算法, 对于不同图形的水印, 由于无法确定检测阈值, 因此, 是不可行的。

在上述可行性分析的基础上, 本文提出了一个可行的鲁棒的矢量地图水印算法, 该算法是一个非盲检测算法, 其特点如下: (1) 对矢量地图的特征点的 DFT 系数幅度和相位同时嵌入水印, 不仅使算法具有一定的抗简化性和鲁棒性, 也减少了算法运算量; (2) 可以以提取的水印与原始水印的相关系数作为检测值, 而对不同的矢量地图图形, 则可用趋于 1 的检测阈值来进行检测; (3) 水印嵌入长度对水印检测影响不大; (4) 不用考虑水印嵌入域是否较好地满足高斯分布; (5) 由于不要求水印序列的均值为零, 因此, 可通过给定的初值和阈值 (如: 0.5) 来生成不同长度的二值水印, 用于嵌入矢量地图中的不同图形; (6) 在保证水印透明性的条件下, 对幅度和相位水印, 水印强度参数 λ 可取固定值; (7) 本文算法对其他类型矢量地图及矢量图形的版权保护也有一定的指导意义或基本可以直接使用。笔者下一步的工作目标是: 寻求实用鲁棒的矢量地图水印盲检测算法。

参考文献 (References)

- 1 XIANG Hui. Review of Digital Watermarking of Graphical Data[J]. Journal of System Simulation, 2002, 14(12): 1649 ~ 1651. [向辉. 图形数据数字水印技术[J]. 系统仿真学报, 2002, 14(12): 1649 ~ 1651.]
- 2 Solachidis V, Nikolaidis N, Pitas I. Watermarking polygonal lines using Fourier descriptors[A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP' 2000) [C], Istanbul, Turkey, 2000: 1955 ~ 1958.
- 3 Ryutarou Ohbuchi, Hiroshi Masuda, Masaki Aono. A shape-preserving data embedding algorithm for NURBS curves and surfaces [A]. In: Proceedings of Computer Graphics International '99 [C], Canmore, Canada, 1999: 170 ~ 177.
- 4 Ryutarou Ohbuchi, Hiroo Ueda, Shuh Endoh. Watermarking 2D

- vector maps in the mesh-spectral domain [A]. In: The fifth International Conference on Shape Modelling and Applications, SMI2003[C], Seoul, Korea, 2003: 216 ~ 228.
- 5 Ryutarou Ohbuchi, Hiro Ueda, Shu Endoh. Robust watermarking of vector digital maps [A]. In: The IEEE International Conference on Multimedia and Expo 2002 [C], Lausanne, Switzerland, 2002: 1 ~ 4.
- 6 Kitamura I, Kanai S, Kishinami T. Watermarking vector digital map using wavelet transformation [A]. In: Proceedings of Annual Conference of the Geographical Information Systems Association (GISA) [C], Tokyo, Japan, 2000: 417 ~ 421.
- 7 Kurihara M, Komatsu N, Arita H. Watermarking vector digital maps [R]. 009-1, Computer Security Information Processing Society of Japan (IPSI), Tokyo, Japan, 2000: 25 ~ 31.
- 8 CHEN Binghe. Random signal processing [M]. Beijing: National Defence Industry Press, 1996. [陈炳和. 随机信号处理 [M]. 北京: 国防工业出版社, 1996.]
- 9 Barni M, Bartolini F, Rosa De, *et al.* Capacity of the watermark-channel: how many bits can be hidden within a digital image? [A]. In: Proceedings of SPIE, Security and Watermarking of Multimedia Contents [C], San Jose, CA, USA, 1999: 437 ~ 448.
- 10 Jain Anil K. Fundamentals of digital image processing [M]. Englewood Cliffs, NJ, USA: Prentice Hall, 1986.
- 11 WU Lun, LIU Yu, ZHANG Jing, *et al.* Geographical information systems: principles and applications [M]. Beijing: Science Press, 2002. [邬伦, 刘瑜, 张晶等. 地理信息系统——原理方法和应用 [M]. 北京: 科学出版社, 2002.]
- 12 Cox Ingemar J, Miller Matthew L, Bloom Jeffrey A. Digital watermarking [M]. Beijing: Publishing House of Electronics Industry, 2003. [[美] Cox Ingemar J, Miller Matthew L, Bloom Jeffrey A. Translate by WANG Ying, HUANG Zhi-be. 数字水印 [M]. 北京: 电子工业出版社, 2003.]
- 13 LI Yu-qi, LIU Hai-yan, ZHAO Lian-wen. Non-parametric statistical method [M]. Chengdu: Publishing House of Southwest Jiaotong University, 1998. [李裕奇, 刘海燕, 赵联文. 非参数统计方法 [M]. 成都: 西南交通大学出版社, 1998.]
- 14 GUO Qing-sheng. Automated map generalization: theoretics and methods [M]. Beijing: Publishing House of Mapping, 2002. [郭庆胜. 地图自动综合理论与方法 [M]. 北京: 测绘出版社, 2002.]
- 15 WANG Tao, LIU Wen-yin, SUN Jia-guang. Using Fourier descriptors to recognize object's shape [J]. Journal of Computer Research and Development, 2002, 39(12): 1714 ~ 1719. [王涛, 赵文印, 孙家广. 傅立叶描述子识别物体的形状 [J]. 计算机研究与发展, 2002, 39(12): 1714 ~ 1719.]
- 16 Weifu H J. Translate by WANG Zhong-de, ZHANG Hui. Theory of discrete and continuous Fourier analysis [M]. Beijing: Publishing House of Beijing University of Posts and Telecommunications, 1991. [[美] H. J. 威佛著. 王忠德, 张惠译离散和连续傅里叶分析理论 [M]. 北京: 北京邮电大学出版社, 1991.]
- 17 LIU Tong, QIU Zheng-ding. The reliability analysis of correlation-based digital watermarking detection [J]. Acta Electronica Sinica, 2002, 30(5): 685 ~ 688. [刘彤, 裘正定. 数字水印相关检测的可靠性研究 [J]. 电子学报, 2002, 30(5): 685 ~ 688.]