

面向二维工程 CAD 图的拓扑完整性校验算法

周 浪¹⁾, 苏智勇^{2)*}, 唐卫清³⁾

¹⁾ (南京财经大学信息工程学院 南京 210023)

²⁾ (南京理工大学自动化学院 南京 210094)

³⁾ (中国科学院计算技术研究所 北京 100190)

(suzhiyong@njust.edu.cn)

摘 要: 拓扑信息是工程 CAD 图的核心, 其完整性直接影响着工程设计的效率和质量. 针对二维异构工程 CAD 图的拓扑完整性认证问题, 提出一种基于半脆弱数字水印技术的通用拓扑校验算法. 首先根据对象间的连接关系构造拓扑图, 并对各对象进行统一编码, 实现异构图纸拓扑关系的统一描述; 然后根据拓扑图建立对应的拉普拉斯矩阵, 利用拉普拉斯谱计算各对象的拓扑特征值; 最后将拓扑特征值嵌入到各对象的几何不变量中. 该算法通过比较提取的特征值序列与对象当前的拓扑特征值来验证其拓扑完整性. 理论分析和实验结果表明, 文中算法对于全局和局部相似性变换、拷贝、镜像和文件格式转换等非恶意操作具有较强的鲁棒性, 同时对于各种恶意拓扑攻击具有精确的定位能力, 可以广泛应用于各种工程 CAD 领域.

关键词: 工程 CAD 图; 拓扑认证; 拓扑完整性; 拉普拉斯矩阵; 数字水印
中图法分类号: TP391.41

Topology Integrity Verification for 2D Engineering CAD Drawings

Zhou Lang¹⁾, Su Zhiyong^{2)*}, and Tang Weiqing³⁾

¹⁾ (College of Information Engineering, Nanjing University of Finance & Economics, Nanjing 210023)

²⁾ (School of Automation, Nanjing University of Science & Technology, Nanjing 210094)

³⁾ (Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

Abstract: Topology information plays the most important role in engineering CAD drawings. The efficiency and quality of engineering design often depend on the topology integrity of drawings. This paper proposes a general topology authentication algorithm, which is inspired by the semi-fragile watermarking scheme, to verify the topology integrity of heterogeneous engineering CAD drawings. Topology graph is firstly built according to the topology relation among objects with uniform codes. Then, the Laplacian matrix is constructed and its Laplacian spectrum is computed to generate topology feature values. These values are then embedded into the geometric invariants of objects. Topology integrity verification is achieved through comparing the extracted binary values with the topology feature values computed according to the current topology relation. Theoretical analysis and experiments have demonstrated that the proposed scheme can not only detect and locate malicious topological modifications, but also yield strong robustness against various topology-preserving modifications, such as global and local similarity transformations, copying, mirroring and file format converting. It's believed that the proposed scheme is a general algorithm and applicable to various kinds of heterogeneous engineering CAD drawings in industry practices.

收稿日期: 2016-05-10; 修回日期: 2016-11-14. 基金项目: 国家自然科学基金(61300160). 周 浪(1982—), 女, 博士, 讲师, 主要研究方向为计算机图形学、模式识别; 苏智勇(1981—), 男, 博士, 副教授, 硕士生导师, CCF 高级会员, 论文通讯作者, 主要研究方向为计算机图形学、计算机视觉; 唐卫清(1965—), 男, 博士, 研究员, 博士生导师, CCF 高级会员, 主要研究方向为 CAD&CG.

Key words: engineering CAD drawing; topology authentication; topology integrity; Laplacian matrix; digital watermarking

图纸是工程 CAD 设计的重要成果,其内容的完整性直接影响着设计审核和工程施工。根据设计对象的不同, CAD 设计可以分为产品 CAD 和工程 CAD。产品 CAD(如机械 CAD)偏重产品细节,侧重造型技术及几何处理算法。工程 CAD(如工厂 CAD),又称 AEC(architecture, engineering, construction),偏重设计对象的宏观结构和拓扑信息描述^[1-2]。为了提高设计效率,工程 CAD 的设计模式已经由传统的串行或者半并行模式向多设计团队、多专业参与,跨地域、分布式的协同设计模式转变^[3]。由于工程 CAD 涉及专业面广,参与人员多,有时还会涉及不同单位之间的合作,所以,工程 CAD 图在设计和传输过程中容易被有意或无意地修改,导致其内容完整性难以鉴别^[4-5]。因此,开展针对工程 CAD 图内容完整性认证的研究,对于防止图纸被恶意篡改或破坏,保障工程 CAD 设计、校审、出图和施工工作的顺利进行十分重要和日益迫切^[6]。

拓扑信息是工程 CAD 图的核心和重点。工程 CAD 图主要由各种对象根据一定的约束关系组合而成;其内容主要包含几何信息、工程属性和拓扑信息。几何信息指对象的尺寸、图形、空间位置等;图形主要由直线、线段、圆、圆弧和面等基本图元组成;工程 CAD 图的绘制需要遵循和符合行业及相关国家标准的规范要求。为了规范设计和提高效率,工程 CAD 设计软件普遍建立了专用图形库。工程属性主要包括工程约束和材料等信息,拓扑信息用来描述对象间的连接关系。在工程 CAD 领域,拓扑信息至关重要且最复杂,是模型的核心所在^[7-9]。设计人员在建模时,将大量设计约束隐含于对象的连接与匹配之中。以工厂 CAD 为例,不仅要描述构件间的连接关系,还要描述具体的连接端口。只有满足管径、端面类型等约束的端口才能相互连接;这些约束是自动生成各类工程图表的基础。因此,对象间拓扑关系的建立与维护是建模的关键问题^[10]。这和面向机械设计的传统 CAD 建模有着明显的不同。所以,拓扑信息应该是工程 CAD 图内容完整性认证的主要对象。

然而,到目前为止,对于二维工程 CAD 图拓扑信息完整性认证问题的关注较少^[6]。现有针对 CAD 图完整性认证和版权保护的工作主要集中于

产品 CAD 领域,研究重点集中在 CAD 图的几何信息方面。解决这类问题主要有 2 种技术途径:数字水印技术和感知哈希技术。数字水印技术通过在待保护对象中嵌入水印信息,实现对象的版权保护和内容完整性认证^[11-16]。其中,半脆弱数字水印技术不仅能抵抗非恶意操作以达到版权保护的目,还能识别和定位各种恶意修改操作、验证内容的完整性。感知哈希技术通过提取表征对象内容的特征产生哈希码,并利用安全信道将该哈希码传输到接收端;接收方通过同样的方法提取接收信息的哈希码,并同接收的哈希码进行比较,确定内容的完整性^[17]。这 2 种技术近年在图像、视频、音频和图形等领域的应用中取得了较大进展^[18-20]。

1) 二维 CAD 图几何信息版权保护

文献[12]提出一种基于 k -means++ 的 CAD 图鲁棒水印算法,它利用 k -means++ 算法将目标物体分组,然后将水印信息嵌入到每组的几何分布之中。文献[13]提出一种基于实体替换与小波均值的二维工程图盲水印算法,通过将图中的直线实体替换为多段线,对变换后的多段线顶点坐标进行二维小波变换,并在小波系数中嵌入水印。文献[14]提出一种基于差值扩大法的二维工程图可逆水印算法,其通过修改线段端点的坐标来嵌入水印信息。Lee 等^[17]针对 CAD 图矢量数据提出一种哈希算法,它利用随机投影法对图中折线的曲率进行投影来提取特征值,并利用二值化方法生成最终的哈希码。

2) 二维 CAD 图几何信息完整性认证

文献[15]提出一种面向工程图版权保护和修改定位的半脆弱水印算法,通过将水印嵌入具有全局特性的属性比值中,并对水印和图元对进行关联处理,实现局部攻击的识别和定位。Peng 等^[11]提出 2 种基于直方图平移的可逆数字水印算法,利用邻接坐标或相对相位的相关性建立差分直方图,基于直方图平移嵌入水印信息,实现几何信息完整性认证。Wang 等^[16]提出一种针对二维向量图的可逆脆弱水印算法,其具有较好的透明性和安全性,但是对于攻击操作的定位能力有待提高。

3) 二维工程 CAD 图拓扑信息完整性认证

文献[6]提出一种针对二维等轴测投影图(isometric drawing, ISO)拓扑完整性认证的数字

水印算法,它利用实体句柄生成各对象的拓扑连接信息和几何不变量,并将这些信息作为水印嵌入到几何不变量中,实现图纸拓扑信息的完整性认证;ISO图是工厂CAD设计中的一种二维图纸。

本文在基于数字水印技术的二维CAD图几何信息版权保护和完整性认证工作的启发之下,提出一种基于半脆弱数字水印技术的二维工程CAD图拓扑完整性认证算法。本文算法根据对象间的拓扑连接关系构造拓扑结构图,并对对象进行编码,实现异构图纸的统一抽象描述;通过建立拓扑结构图的拉普拉斯矩阵,实现对象拓扑特征值的计算和提取。本文算法利用数字水印技术,将对象的拓扑特征值嵌入到该对象的几何不变量中,实现图纸拓扑关系的完整性认证。

1 拓扑信息完整性校验算法

1.1 算法概述

本文算法分为2部分:拓扑特征值嵌入和拓扑信息校验。对于工程CAD图中的每个待保护对象,首先提取它的拓扑特征值并作为水印信息;然后构造该对象的几何不变量作为水印载体;最后将特征值嵌入载体中实现对象拓扑关系的保护。在校验时,首先从对象几何信息中提取水印信息;然后根据对象当前拓扑信息计算拓扑特征值;最后通过对比这2种特征值实现对象拓扑信息完整性的校验。

待保护对象通常是指组成工程CAD图的所有几何对象。由于本文拓扑完整性校验是通过在对象几何信息中嵌入拓扑特征值实现的,因此,该对象需要具备一定的几何复杂度(如包含一定数量的顶点),以便于构造几何不变量作为水印载体。

1.2 拓扑特征提取

拓扑特征提取的目的是为了将待保护对象与其连接对象之间的拓扑关系映射为唯一的特征值(也称为拓扑水印)。该特征值取决于对象类型及其之间的连接关系,其对于各种保拓扑操作(在保证对象间拓扑关系不变的前提下,对对象几何信息进行的各种编辑操作)具备强鲁棒性。

1.2.1 对象编码

对象编码是赋予每种对象一个唯一的编码值。在工程CAD领域,由于设计阶段和标准的不同,同一对象可能存在多种不同的图形描述方式。为了解决图形表达的异构性,本文首先对工程CAD图的

各类对象进行统一编码。由于工程CAD软件配有专门的图形库,因此,可以根据图形库和各行各业的专业知识,对所有对象进行分类,再统一编码。以流程工厂CAD为例,首先将对象分为设备、元件和建筑构件3大类,然后再对每一大类进行细分。例如,元件可以细分为管子、管件、阀门等。最后,根据分类级别数给每种对象设计一个唯一的 m_c 位编码值。

1.2.2 拓扑图构造

拓扑图用来描述对象间的连接关系,是一种无向加权图。在工程CAD领域,不同CAD软件所采用的拓扑描述方式可能不同。常见拓扑描述方法主要有2种:基于连接点和基于对象存储顺序的拓扑关系描述^[8]。这些描述方法或依赖对象句柄建立对象间的拓扑关系,或利用对象间的存储顺序来维护它们之间的连接关系。为了提高算法通用性,本文根据各类工程CAD图采用的具体拓扑建模方式获取对象间的拓扑关系,并建立统一的无向加权图 G 。

拓扑图 $G(V, E)$ 是一个点集为 $V = \{v_1, v_2, \dots, v_n\}$,边集为 $E = \{(v_i, v_j) | v_i, v_j \in V\}$ 的无向加权图。其中, v_i 表示工程CAD图中的待保护对象, $d(v_i)$ 为点 v_i 的度,边 (v_i, v_j) 表示对象 v_i 和 v_j 相连,其权值记为

$$w_{ij} = C_i \times 10^{-m_c} + C_j \times 10^{-2m_c} \quad (1)$$

其中, C_i 和 C_j 分别为对象 v_i 和 v_j 的编码值; m_c 为对象编码的位数。为了保证权值 w_{ij} 的唯一性,式(1)中要求 $C_i \leq C_j$ 。本文利用式(1)将具有连接关系的对象进行编码,以量化的形式描述对象间的拓扑连接关系,从而克服不同图纸在几何和拓扑表达上的差异性。

1.2.3 拓扑特征值计算

对于每个顶点 v_i ,首先构造其子图 $G_{v_i}(V_{v_i}, E_{v_i})$, $V_{v_i} = \{v_i\} \cup \{v_j | (v_i, v_j) \in E\}$, $E_{v_i} = \{(v_i, v_j) | v_j \neq v_i \in V_{v_i}\}$;然后建立图 G_{v_i} 的拉普拉斯矩阵 $L(G_{v_i})$,利用 $L(G_{v_i})$ 的拉普拉斯谱计算对象 v_i 的拓扑特征值。

对于具有 k 个顶点的子图 G_{v_i} ,其拉普拉斯矩阵定义为 $L(G_{v_i}) = (l_{ij})_{k \times k} = D(G_{v_i}) - A(G_{v_i})$,其中 $D(G_{v_i}) = (d_{ij})_{k \times k}$ 为图 G_{v_i} 的度对角矩阵, $A(G_{v_i}) = (a_{ij})_{k \times k}$ 为图 G_{v_i} 的邻接矩阵;矩阵元素分别为

$$d_{ij} = \begin{cases} d(v_i), & i = j \\ 0, & i \neq j \end{cases}$$

$$a_{ij} = \begin{cases} w_{ij}, & (v_i, v_j) \in E \\ 0, & (v_i, v_j) \notin E \end{cases}$$

那么矩阵 $L(G_{v_i})_{k \times k}$ 的元素 l_{ij} 可以表示为

$$l_{ij} = \begin{cases} d(v_i), & i = j \\ -w_{ij}, & i \neq j \wedge (v_i, v_j) \in E \\ 0, & \text{else} \end{cases}$$

为了保证矩阵 $L(G_{v_i})$ 的唯一性, 本文根据图 G_{v_i} 中各对象编码值的大小, 将它们按照升序排列; 然后再建立对应的度对角矩阵 $D(G_{v_i})$ 和邻接矩阵 $A(G_{v_i})$. 根据拉普拉斯矩阵的性质可知, $L(G_{v_i})$ 是个 k 阶的对称半正定矩阵, 具有 k 个非负的实数特征值^[21]. 不失一般性, 假设它们按照从大到小的顺序排列为 $\lambda_{k-1}(G_{v_i}) \geq \lambda_{k-2}(G_{v_i}) \geq \dots \geq \lambda_0(G_{v_i})$; 矩阵 $L(G_{v_i})$ 的所有特征值也称为图 G_{v_i} 的拉普拉斯谱.

在求取图 G_{v_i} 的拉普拉斯谱之后, 本文利用

$$\{w_i\}_{n_w} = F_{sel}(\lambda_{k-1}(G_{v_i}), n_w, K) \oplus \dots \oplus F_{sel}(\lambda_0(G_{v_i}), n_w, K) \quad (2)$$

提取 n_w 位的二进制序列 $\{w_i\}_{n_w}$, 将其作为对象 v_i 的拓扑特征值. 其中, 函数 $F_{sel}(f, n_w, K)$ 实现在密钥 K 的控制下, 从浮点数 f (用 IEEE-754 双精度浮点数格式表示) 的尾数位中提取 n_w 位的二进制序列; \oplus 为异或运算.

如图 1 所示, 本文以一幅简单的工程 CAD 图为例, 说明对象 v_4 (专业名称为“三通”) 拓扑特征值的计算过程. 图中 7 个对象 (v_1, \dots, v_7) 通过不同的连接线(管子)相连. 首先根据对象间的连接关系构造拓扑结构图 G 和对象 v_4 对应的子图 G_{v_4} , 然后建立 G_{v_4} 的拉普拉斯矩阵, 最后根据其拉普拉斯谱生成对应的拓扑特征值.

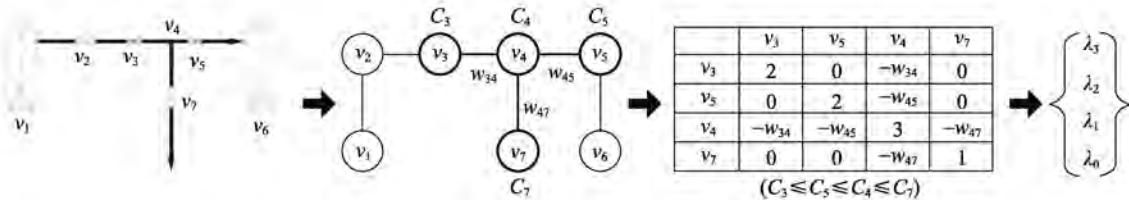


图 1 对象拓扑特征值计算示意图

1.3 载体构造

本文根据待保护对象的几何信息选择和构造几何不变量, 并将其作为拓扑特征值的嵌入载体. 几何不变量对于非恶意操作必须具备较强的鲁棒性, 能够保证嵌入信息的完整性和安全性. 常用几何不变量主要有角度比值、距离比值和面积比值. 根据待保护对象特点, 本文选择距离比值嵌入水印信息.

首先, 利用对象顶点数据, 按照下列规则构造和选择合适的三角形. 假设组成对象的所有二维顶点数据的集合为 $P = \{p_1, p_2, \dots, p_n\}$; 其中, $p_i = \{p_{ix}, p_{iy}\}$, n 为顶点数量. 基于顶点集合 P , 构造所有可能的三角形集合 $T = \{t_1, t_2, \dots, t_m\}$ (m 为三角形数量), 且 $A_i \geq A_{i+1}$; 其中 A_i 表示 Δt_i 的面积. 本文从三角形集合 T 中选择最大面积的 $\Delta t_1 (\Delta p_a p_b p_c)$ 作为嵌入对象.

然后, 对于 $\Delta p_a p_b p_c$, 设 $P' = P - \{p_a, p_b, p_c\}$, p^c 为 P' 的几何中心, $p^c = \frac{1}{|P'|} \sum_{p_i \in P'} p_i$. $|p^c p_a|$,

$|p^c p_b|$, $|p^c p_c|$ 分别表示 p^c 到 $\Delta p_a p_b p_c$ 3 个顶点的距离, 且假设

$$|p^c p_a| \geq |p^c p_b| \geq |p^c p_c| \quad (3)$$

$R_{p_a p_b p_c} = \max\{|p_i p^c|\}$, $p_i \in P'$ 为 P' 中各点到 p^c 距离的最大值. 本文选择距离比值 $r = |p^c p_a| / R_{p_a p_b p_c}$ 作为水印嵌入载体, 如图 2 所示. 其中, p_a 称为水印嵌入点, p_b 和 p_c 称为 $\Delta p_a p_b p_c$ 的标记点. 由于距离比值具有平移、旋转和缩放不变性, 因此在距离比值中嵌入拓扑特征值对于非恶意操作具有较强的鲁棒性.

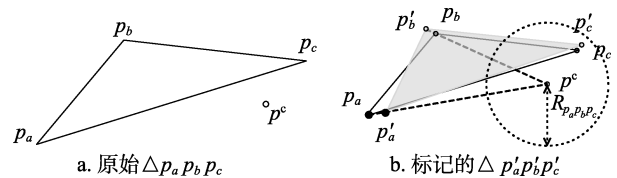


图 2 载体构造和标记示意图

1.4 特征值嵌入

本文通过将拓扑特征值嵌入待保护对象的几何不变量中, 实现对象拓扑连接关系的保护和

验证. 在 CAD 图中, 顶点坐标等属性值通常使用 IEEE-754 定义的双精度浮点数格式表示. 双精度浮点数可以表示高达 16 位左右的有效数字, 而实际应用中通常只需要 6 位有效数字^[15]; 这给水印信息的嵌入提供了一定的空间. 设 $\{w_i\}_{n_w}$ 为提取的拓扑特征值即水印信息, $r = |p^c p_a| / R_{p_a p_b p_c}$ 为构造的几何不变量, 掩码 B_{mask} 是一组 n_w 位二进制序列; 其中, 0 代表无效位, 1 代表有效位. 算法通过位替换方法按照一定规则从 r 的尾数部分选择 n_w 位进行替换, 具体过程如下:

Step1. 首先从 r 的双精度浮点数表示中提取尾数部分.

Step2. 按照由高位到低位的顺序, 将二进制水印序列 $\{w_i\}_{n_w}$ 的每一位和掩码 B_{mask} 中对应的有效位进行与运算, 并将结果写入到该有效位对应的尾数位中, 得到嵌入水印后的值 $r' = \frac{|p^c p'_a|}{R_{p_a p_b p_c}}$, 如图 3a 所示.

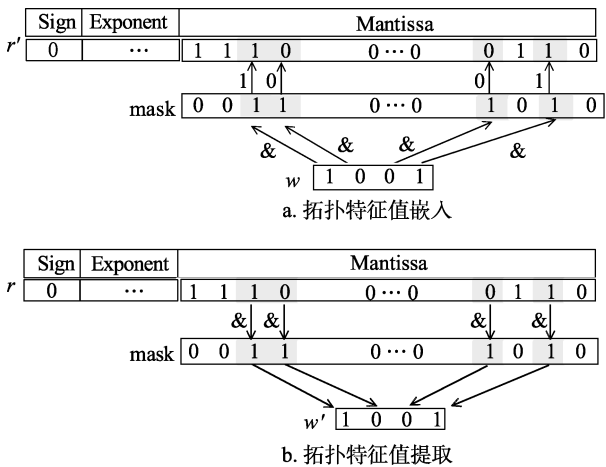


图 3 拓扑特征值嵌入和提取示意图

Step3. 得到 r' 之后, 算法通过调整水印嵌入点 p_a 的坐标, 实现拓扑特征值的嵌入; p'_a 为 p_a 调整之后的坐标. 算法通过沿着矢量 $p^c p_a$ 的方向移动点 p_a , 使其满足 $p'_a = p^c + \frac{p^c p_a}{|p^c p_c|} \times |p^c p'_a| = p^c + \frac{p^c p_a}{|p^c p_c|} \times r' \times R_{p_a p_b p_c}$, 从而实现水印信息的嵌入.

1.5 载体标记

为了在拓扑校验阶段, 准确找到嵌入了拓扑水印的三角形, 需要对 $\triangle p_a p_b p_c$ 进行标记, 使其满足特定的约束条件.

假设 $r_1 = |p^c p_b| / R_{p_a p_b p_c}$, $r_2 = |p^c p_c| / R_{p_a p_b p_c}$, 本文通过调整标记点 p_b 和 p_c 的坐标, 使 $\triangle p_a p_b p_c$ 同时满足约束条件

$$\begin{cases} r'_1 = \frac{|p^c p'_b|}{R_{p_a p_b p_c}} = \left[r_1 / \delta_{p_a p_b p_c} \right] \times \delta_{p_a p_b p_c} + n_1 \times \delta_{p_a p_b p_c} \\ r'_2 = \frac{|p^c p'_c|}{R_{p_a p_b p_c}} = \left[r_2 / \delta_{p_a p_b p_c} \right] \times \delta_{p_a p_b p_c} + n_2 \times \delta_{p_a p_b p_c} \\ \xi = \delta_{p_a p_b p_c} \times R_{p_a p_b p_c} \end{cases} \quad (4)$$

从而实现标记操作. 其中, n_1 和 $n_2 \in \mathbb{N}$; $\delta_{p_a p_b p_c}$ 为 $\triangle p_a p_b p_c$ 的量化步长; ξ 为全局误差控制量; p'_b 和 p'_c 为调整之后的坐标, 分别为

$$\begin{cases} p'_b = p^c + \frac{p^c p_b}{|p^c p_b|} \times |p^c p'_b| = p^c + \frac{p^c p_b}{|p^c p_b|} \times r'_1 \times R_{p_a p_b p_c} \\ p'_c = p^c + \frac{p^c p_c}{|p^c p_c|} \times |p^c p'_c| = p^c + \frac{p^c p_c}{|p^c p_c|} \times r'_2 \times R_{p_a p_b p_c} \end{cases} \quad (5)$$

算法采取分别沿着矢量 $p^c p_b$ 和 $p^c p_c$ 的方向调整 p_b 和 p_c 坐标的方式求解式(4). 可以看出, 随着 n_1 和 n_2 的取值不同, 满足上述条件的点 p'_b 和 p'_c 也不唯一. 假设满足式(5)的 p'_b 和 p'_c 的解的集合分别为 P_b 和 P_c ; 为了降低标记操作导致的几何形变, 可以通过如下方式求解 p'_b 和 p'_c 的值, 即

$$\begin{cases} p'_b = \arg \min_{p'_b \in P_b} \left\| |p^c p_b| - |p^c p'_b| \right\|, \text{ s.t. } |p^c p'_b| < |p^c p_a| \\ p'_c = \arg \min_{p'_c \in P_c} \left\| |p^c p_c| - |p^c p'_c| \right\|, \text{ s.t. } |p^c p'_c| < |p^c p_a| \end{cases}$$

标记后的 $\triangle p_a p_b p_c$ 满足 r_1 和 r_2 能被 $\delta_{p_a p_b p_c}$ 整除, 同时保持 $|p^c p_a| > |p^c p_b| \geq |p^c p_c|$. 在拓扑校验阶段, 可以通过这些约束条件寻找到被嵌入了拓扑特征值的三角形载体.

图 2 所示为载体构造和标记示意图. 图 2a 中的 $\triangle p_a p_b p_c$ 为初始三角形, p^c 为 P' 的几何中心. $\triangle p'_a p'_b p'_c$ 为嵌入了拓扑特征值并标记之后的三角形, p'_a 为嵌入特征值后的坐标, p'_b 和 p'_c 为标记后的坐标点.

1.6 拓扑认证

对于已经嵌入拓扑特征值的工程 CAD 图, 算法通过逐一校验其保护对象的拓扑完整性, 实现整幅工程 CAD 图拓扑完整性的认证. 主要步骤如下:

Step1. 根据图纸中各对象的当前拓扑关系建立图纸的拓扑结构图, 并对图中每个保护对象执行 Step2~Step6.

Step2. 按照第 1.3 节的方法, 根据组成该对象的几何顶点信息构造可能的三角形集合, 并按照面积从大到小排列. 由于特征值嵌入和载体标记操作可能会导致三角形面积发生变化, 而三角形集合中面积最大的三角形不一定是之前嵌入了拓扑特征值的三角形; 因此, 需要

依次从集合中选择三角形, 找到同时满足式(3)(4)约束条件的标记 $\triangle p_a p_b p_c$.

Step3. 计算 $\triangle p_a p_b p_c$ 对应的几何不变量 $r = |p^c p_a| / R_{p_a p_b p_c}$.

Step4. 利用掩码 B_{mask} , 从 r 的尾数中提取嵌入的水印信息 w' , 如图 3b 所示.

Step5. 构造该对象对应的拓扑子图, 利用式(2)计算其对应的拓扑特征值 w .

Step6. 采用归一化相关系数(normalized correlation, NC)计算提取到的拓扑特征值 w' 和当前拓扑特征值 w 之间的相似度^[11]. 如果 $NC=1$, 说明该对象的拓扑信息没有被修改; 否则, 表明该对象的拓扑关系发生变化. NC 的计算公式为

$$NC = \frac{\sum_{i=0}^{n_w-1} w_i \times w'_i}{\sqrt{\sum_{i=0}^{n_w-1} (w_i)^2} \times \sqrt{\sum_{i=0}^{n_w-1} (w'_i)^2}} \quad (6)$$

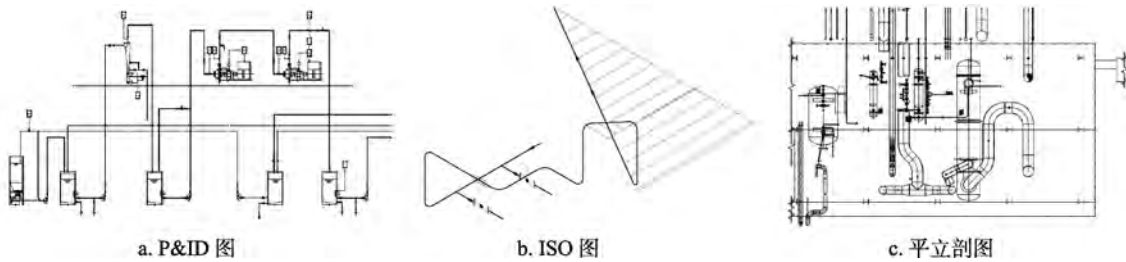


图 4 工程 CAD 图示例

$$\begin{cases} P_R = \frac{M_x}{M} \\ T_L = \frac{M_c}{M_t} \end{cases} \quad (7)$$

其中, M 为图纸中总的保护对象数; M_x 为嵌入拓扑特征值能够被正确提取的总的对象数; M_t 为被恶意攻击的对象总数; M_c 为正确检测出的被攻击对象总数. 这里所指的嵌入拓扑特征值能够被正确提取是指提取的特征值和当前计算的特征值满足式(6)约束. 可以看出, 比值 P_R 越大, 说明算法的拓扑校验能力越强; T_L 值越大, 算法的攻击定位能力越强.

本文讨论的图形编辑操作分为恶意操作和非恶意操作 2 大类. 恶意操作是指那些会改变对象间拓扑连接关系的操作, 例如, 删除和增加对象、断开连接等操作. 非恶意操作是指在保证对象间拓扑关系不变的前提下所进行的操作, 例如, 全局和局部相似性变换(旋转、缩放、平移)、拷贝、镜像、文件格式转换等. 需要指出的是, 由于工程 CAD

2 实验与分析

为了验证算法的有效性, 本文以工厂 CAD 中的 ISO 图、管道仪表流程图(piping and instrumentation diagram, P&ID)和平立剖图 3 种代表性的异构二维工程 CAD 图为例, 各选择 10 幅图进行实验. 图 4 所示为 3 种图纸的示意图, 图纸的数值精度为 10^{-3} . 根据图形库中各对象的种类数将对象的编码位数设 $m_c=4$.

本文使用 RMSE(root mean square error)^[11]来评估拓扑特征值的嵌入所引起的图形几何形变, 即 $RMSE = \frac{1}{N} \|V - V'\|$. 其中, V 和 V' 分别表示原始工程 CAD 图和嵌入了拓扑特征值的工程 CAD 图中相对应的顶点; N 为总的顶点个数. 此外, 本文分别使用 P_R 和 T_L 来衡量算法的鲁棒性和攻击定位能力, 即

图设计的特殊性, 本文所讨论的图纸编辑操作均是针对图中对象级别的操作, 而不考虑针对基本图元的操作.

2.1 参数讨论

2.1.1 掩码 B_{mask}

掩码 B_{mask} 的取值影响着算法的不可见性、鲁棒性和安全性; 它是一组总长为 52 位的二进制序列, 包含了 n_w 位有效位‘1’.

就不可见性而言, 在 n_w 确定的情况下, 有效位的位置越靠近高位, 则特征值嵌入操作引起的几何形变就会越大. 相反, 若有效位的位置越接近低位, 则引起的几何形变就会越小. 本文以图 4a 所示 P&ID 图为例, 在特征值位数 $n_w=8$ 时, 计算了有效位处在 B_{mask} 中不同位置时对应的几何形变程度, 结果如图 5a 所示. 可以看出, RMSE 值和嵌入位置成一定的比例关系. 随着有效位的位置越接近低位, 算法的不可见性越强.

就鲁棒性而言, 算法利用 B_{mask} 将特征值嵌入

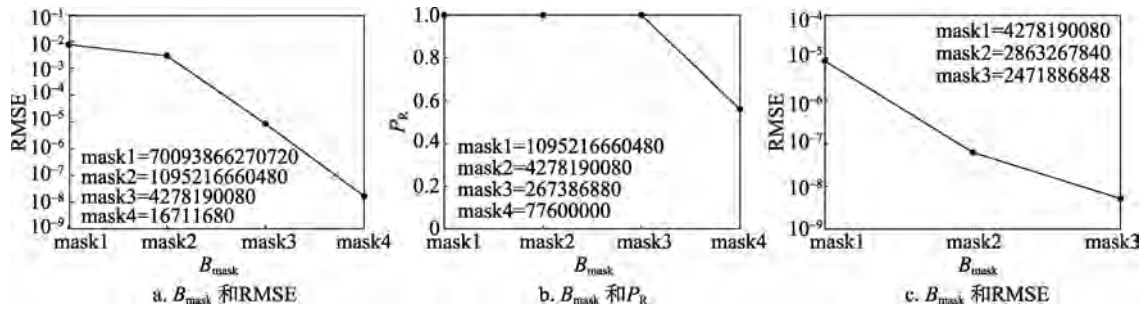


图 5 掩码 B_{mask} 对不可见性和鲁棒性的影响($n_w = 8$)

到图纸中, 当对图纸进行诸如旋转、缩放和平移等非恶意操作之后, 理论上算法应该仍然能够准确地从图纸中提取出相同的特征值. 然而, 在实际应用中, 如果 B_{mask} 的取值不当, 则容易导致算法在拓扑校验阶段发生误判. 这是由于浮点数运算的不精确性所引起的^[22]. 本文以图 4c 所示平立剖图为例, 在 $n_w = 8$ 时, 采用不同的 B_{mask} 值, 首先实现拓扑特征值的嵌入, 然后再对图纸进行上述各种非恶意操作; 最后, 计算校验阶段能够正确提取嵌入值的比例, 结果如图 5b 所示. 可以看出, 当有效位的最高位低于 28 时, 算法的误检率开始上升; 当有效位的最高位高于 28 时, 算法能够正确地从提取嵌入的拓扑特征值, 实现拓扑完整性的准确校验.

就安全性而言, 掩码 B_{mask} 的另外一个重要用途是用作算法的密钥. 例如, 当 $n_w = 6$ 时, 理论上 B_{mask} 的组合可以有 $C_{52}^6 = 20358520$ 种可能性; 并且, 随着有效位 n_w 的增大, B_{mask} 的组合数也会随之增加. 因此, 攻击者在不知道掩码 B_{mask} 的前提下, 即使掌握了拓扑特征值的嵌入和提取方法, 也很难篡改嵌入的拓扑特征值. 本文以图 4b 所示 ISO 图为例, 计算了采用不同 B_{mask} 值时对应的 RMSE 值, 结果如图 5c 所示. 可以看出, 在有效位最高位取值相同的情况下, 可以通过适当改变其他有效位的位置分布来提高算法的安全性, 且不会导致图纸几何形变的增加.

根据上述分析, 本文通过大量实验发现, 当有效位位于 [39, 26] 时能够满足实际应用需求. 本文后续实验所用掩码的十进制数为 $B_{\text{mask}} = 50113544192$.

2.1.2 n_w

n_w 表示提取的拓扑特征值位数, 同时, 也决定了掩码 B_{mask} 中有效位的位数. 因此, n_w 的取值影响着算法的不可见性和鲁棒性.

就不可见性而言, n_w 越大, 意味着掩码 B_{mask} 中的有效位就越多, 在同等情况下由于特征值嵌入操作所引起的几何形变也会越大. 就鲁棒性而言, 拓扑特征值的位数越多, 那么在校验阶段, 当对象被恶意攻击或没有嵌入过特征值时, 从该对象中提取的二进制序列与当前计算的特征值发生重合的概率就会越低. 在掩码 B_{mask} 中最高有效位位置相同的情况下, n_w 取不同值时对算法不可见性的影响如图 6 所示. 结合上述对于 B_{mask} 取值的分析和实验, 实验中选择的拓扑特征值位数为 $n_w = 8$.

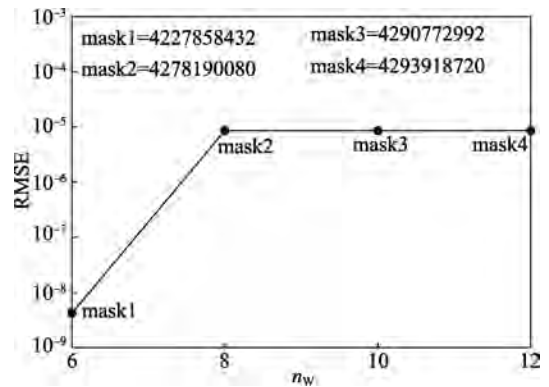


图 6 n_w 和 RMSE 的关系

2.1.3 密钥 K

K 是一组总长 52 位、包含 n_w 位随机分布的有效位的二进制序列, 在函数 $F_{\text{sel}}(f, n_w, K)$ 中, 用于从 f 尾数中对应的位置提取相应的二进制数, 并按照从高位到低位的顺序排列, 从而得到最终的二进制序列, 提高算法的安全性. 实验中所用密钥的取值(十进制)为 $K = 321066123755520$.

2.1.4 ξ

ξ 用于标记已嵌入拓扑特征值的 $\Delta p_a p_b p_c$, 使得 $r_1 = |p^c p_b| / R_{p_a p_b p_c}$ 和 $r_2 = |p^c p_c| / R_{p_a p_b p_c}$ 能被量化步长 $\delta_{p_a p_b p_c}$ 整除. 标记操作是通过调整标记点

p_b 和 p_c 实现的; 根据式(4)可以看出, 坐标调节的增量值是 $\xi = \delta_{p_a p_b p_c} \times R_{p_a p_b p_c}$. 本文采取固定坐标调节增量值 ξ , 根据各对象的实际尺寸计算该对象对应量化步长值的策略, 即 $\delta_{p_a p_b p_c} = \xi / R_{p_a p_b p_c}$. 其中, 增量值 ξ 的大小是根据图纸数据精度的需求确定. ξ 的取值会影响图纸几何形变的程度, 且随着 ξ 的增大, 图纸的几何形变也会随之增大. 图 7 给出了 ξ 在不同取值情况下对于图纸几何形变的影响. 本文实验中, 根据图纸的精度要求, ξ 的取值为 10^{-4} .

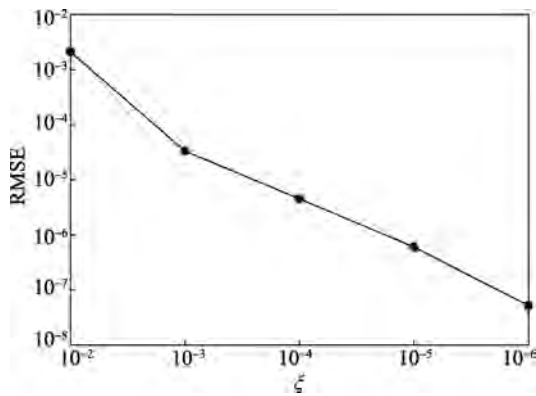


图 7 ξ 和 RMSE 的关系

2.2 定位能力分析

在图纸编辑过程中, 可能导致对象拓扑关系发生变化的操作主要包括增加、删除和替换连接对象 3 种方式. 算法在拓扑校验阶段应该能够发现和定位那些拓扑连接关系发生变化的对象.

2.2.1 增加连接对象

增加连接对象即在图中增加新的对象, 并将它们与已有对象连接. 对原有对象而言, 新增加的连接对象改变了该对象原有的拓扑结构, 使得节点之间的连接关系以及边的权值发生变化, 进而引起对应的拉普拉斯矩阵发生变化, 最终导致根据当前连接关系计算的拓扑特征值与之前嵌入该对象的拓扑特征值之间的差异. 另一方面, 对新增

加的对象而言, 其自身几何信息中并没有被嵌入过拓扑特征值信息. 所以, 在拓扑检验阶段就无法从中提取有效拓扑特征值, 并与对象当前的拓扑特征值进行比较. 因此, 算法可以有效检测并定位出此类攻击.

图 8a 所示为原始 P&ID 图的一部分, $w_{AB} = w_{AD} = 0.12006030, w_{BC} = 0.24006030$, 对象 A, B 的拓扑特征值分别为 1111101, 10000011. 实验中, 本文给对象 B 和 C 所在的管线增加了一个由对象 E 和 F 组成的简单分支管线, 如图 8b 所示; 其中 $w_{BE} = 0.22006030$. 以对象 B 为例, 其连接对象由之前的 A 和 C 变为 A 和 E, 且 C 和 E 类型不同. 因此, 在校验 B 的拓扑完整性时, 其当前拓扑特征值变为 00011010, 和提取出的原始拓扑特征值 10000011 不同. 因此, 可以确定对象 B 的连接关系发生了变化.

2.2.2 删除连接对象

删除连接对象即将待删除对象和与它相连的对象断开, 并从图中删除. 对于原有的连接对象而言, 删除操作会减少其拓扑图的节点个数, 从而改变对应的拉普拉斯矩阵, 最终导致当前拓扑特征值与嵌入的拓扑特征值之间的差异. 因此, 在拓扑检验阶段, 算法同样可以有效检测并定位出此类攻击.

如图 8c 所示, 实验中删除了对象 D 所在的管线. 这导致对象 A 的连接对象由之前的 B 和 D 变为现在的 B. 在校验 A 的拓扑完整性时, 当前的拓扑特征值变为 00100000, 不同于提取出的原始拓扑特征值 1111101. 因此, 可以确定对象 A 的连接关系发生了变化.

2.2.3 替换连接对象

替换连接对象即将原先的连接对象删除, 然后加入并连接新的对象(不同种类). 因此, 该操作方式可以理解为删除和增加操作的组合. 它虽然不改变被攻击对象拓扑图的结构, 但是由于新连接

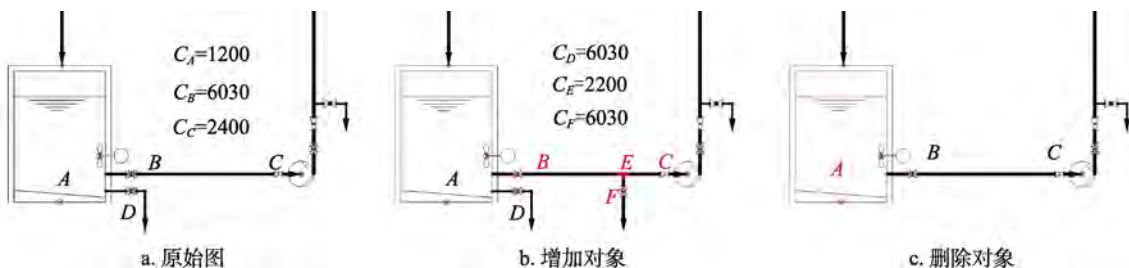


图 8 拓扑攻击定位示意图

对象的编码和原有对象不同, 因此, 最终也会导致拓扑特征值的差异.

为了测试算法对于恶意攻击的定位能力, 本文设计了如下攻击实验: 对于每张图纸, 选择 3 组对象, 其中每组的对象数约占图纸总对象数的 20%. 对于第 1 组, 给各对象分别增加 1 个新的连接对象; 对于第 2 组, 分别从各对象的连接对象中删除一个连接对象; 对于第 3 组, 分别用一个新的不同类别的对象去替换各对象的一个原有连接对象. 最后, 利用本文校验算法对这些图纸进行了拓扑校验, 并根据式(7)计算出了被攻击对象的正确提取率 T_L , 结果如表 1 所示. 可以看出, 在上述给定的参数条件下, 本文校验算法具有较好的攻击定位能力; 它和文献[6]一样, 能够实现对象级局部拓扑攻击的准确定位.

表 1 定位能力测试结果(T_L)

图纸	增加对象	删除对象	替换对象
第 1 组	1	1	1
第 2 组	1	1	1
第 3 组	1	1	1

2.3 鲁棒性分析

本文根据编辑操作是否会改变对象句柄值, 将它们分为 2 类进行讨论. 一类是相似性变换(旋转、缩放、平移), 这类操作不会改变对象的句柄值; 一类是拷贝、镜像和文件格式转换等操作, 这类操作很容易导致对象句柄值的改变甚至丢失^[12]. 句柄是 CAD 图的基本属性, 它能唯一标志图中各个对象, 在许多水印算法中获得了广泛应用^[12,23].

2.3.1 相似性变换

全局和局部相似性变换(旋转、缩放、平移)是图形编辑过程中常见的非恶意操作. 在工程 CAD 中, 为了使图纸布局更加美观, 以及便于信息标注, 在保证对象拓扑关系不变的前提下, 通常允许对图中部分对象进行局部相似性变换, 如图 9 所示. 这类操作会不同程度地改变对象的几何信息, 但不会影响图纸的功能. 因此, 算法对于此类操作需要具备较强的鲁棒性. 下面主要从拓扑特征值计算、载体构造和标记、特征值嵌入 3 个方面分析算法的鲁棒性:

1) 拓扑特征值计算. 根据第 1.2.2, 1.2.3 节的描述, 拓扑图构造和拓扑特征值的计算只和各对象的类型以及对象之间的拓扑连接关系有关. 图纸编辑过程中的非恶意操作并不会引起对象类型

和拓扑关系的变化, 因此拓扑特征值的计算是鲁棒的.

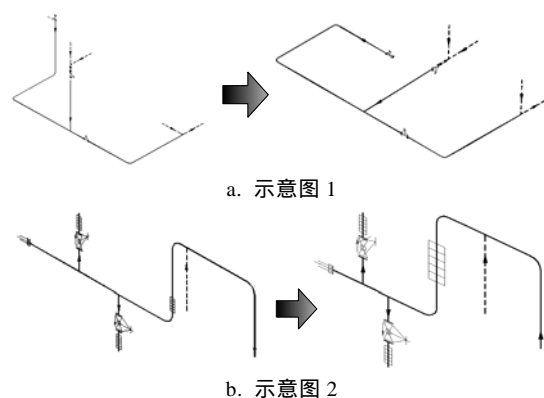


图 9 局部相似变换示意图

2) 特征值嵌入. 为了将拓扑特征值嵌入待保护对象, 本文算法首先根据对象的几何信息构造和选择 $\Delta p_a p_b p_c$; 然后利用距离比值, 即 $r = |p^c p_a| / R_{p_a p_b p_c}$ 作为水印嵌入载体. 由于距离比值具有旋转、缩放和平移不变性, 因此, 本文的特征值嵌入方法对于相似变换也是鲁棒的.

3) 载体标记. 为了对嵌入了拓扑特征值的三角形进行标记, 本文算法同样基于距离比值, 利用量化的方法实现标记, 其中量化步长的取值亦考虑了缩放攻击; 所以, 载体标记方法也是鲁棒的.

综上所述, 此类变换不会改变图中各对象的拓扑关系以及图纸的功能, 也不会影响嵌入的拓扑特征值. 为了验证算法的鲁棒性, 本文对实验用图纸分别进行了全局相似性变换和局部相似性变换操作测试. 对于局部相似性变换操作, 实验主要通过随机选取图纸中的对象, 在保证拓扑关系不变的前提下进行各种旋转、缩放和平移操作; 测试结果如表 2 所示. 可以看出, 在上述给定参数条件下, 本文算法的提取正确率为 100%, 误判率为 0.

表 2 鲁棒性测试结果

测试指标	旋转	缩放	平移
P_R	1	1	1
T_L	0	0	0

2.3.2 拷贝、镜像和文件格式转换

拷贝和镜像操作会生成新的对象, 并保持这些对象间的拓扑关系不变. CAD 系统在生成这些新对象时会自动给它们分配新的句柄值, 用来唯一标志这些对象. 此外, 在实际应用中, 经常需要

进行文件格式转换操作. 由于不同 CAD 系统采用的句柄编码方式可能不同, 也会导致转换后文件中各对象的句柄发生变化. 然而, 由于本文算法的各个环节并不依赖句柄, 因此, 其对于此类攻击是鲁棒的.

2.4 算法比较

到目前为止, 关于二维工程 CAD 图拓扑完整性认证的工作还不多. 本文主要与文献[6]的工作进行对比, 从通用性和鲁棒性 2 个方面分析和总结本文算法的优点:

1) 通用性. 为了解决不同种类工程 CAD 图在几何和拓扑建模方式上的异构性, 提出了一种通用的拓扑特征提取方法. 根据对象类型对所有对象进行了统一编码, 然后建立图纸的拓扑结构图, 实现对异构图纸拓扑信息的统一描述. 算法基于各对象的拓扑子图建立对应的拉普拉斯矩阵, 利用拉普拉斯谱生成各对象的拓扑特征值. 文献[6]主要利用对象的句柄值生成水印信息, 适用范围有限.

2) 鲁棒性. 文献[6]在水印生成、载体构造、水印嵌入和提取等阶段均依赖对象的句柄值, 鲁棒性相对较弱. 本文提出了一种利用对象几何特点构造和选择载体的方法, 且在特征值嵌入和载体标记过程中不依赖句柄值. 因此, 本文算法能够抵抗拷贝、镜像和文件格式转换等此类易导致对象句柄发生变化的常用攻击, 具有更强的鲁棒性.

根据上述分析以及实验证明, 本文算法在取得和文献[6]同等攻击定位能力的同时, 显著提高了算法的通用性和鲁棒性.

3 结 语

本文分析和讨论了二维异构工程 CAD 图的拓扑完整性认证问题, 提出了一种基于半脆弱数字水印技术的通用拓扑认证算法. 算法基于拓扑图和拉普拉斯矩阵实现了异构图纸拓扑信息的统一描述和拓扑特征值的计算, 并根据对象几何特性, 利用距离比值构造和标记嵌入载体, 实现拓扑特征值的嵌入; 本文以工厂 CAD 中的相关图纸为例进行了实验. 理论分析和实验结果表明, 本文算法具备精确的对象级别攻击定位能力, 较强的鲁棒性和实用性. 本文算法可以推广到其他工程 CAD 领域.

参考文献(References):

- [1] Tang Rongxi. CAD/CAM technology[M]. Beijing: Beihang University Press, 1994: 2-8(in Chinese)
(唐荣锡. CAD/CAM 技术[M]. 北京: 北京航空航天大学出版社, 1994: 2-8)
- [2] Farr E R P, Piroozfar P A E, Robinson D. BIM as a generic configurator for facilitation of customisation in the AEC industry[J]. Automation in Construction, 2014, 45: 119-125
- [3] Li W D, Lu W F, Fuh J Y H, *et al.* Collaborative computer-aided design—research and development status[J]. Computer-Aided Design, 2005, 37(9): 931-940
- [4] Kim T, Cera C D, Reglib W C, *et al.* Multi-level modeling and access control for data sharing in collaborative design[J]. Advanced Engineering Informatics, 2006, 20(1): 47-57
- [5] Mun D, Hwang J, Han S. Protection of intellectual property based on a skeleton model in product design collaboration[J]. Computer-Aided Design, 2009, 41(9): 641-648
- [6] Su Z Y, Yang X, Liu G J, *et al.* Topology authentication for piping isometric drawings[J]. Computer-Aided Design, 2015, 66: 33-44
- [7] Su Z Y, Zhou L, Liu G J, *et al.* Authenticating topological integrity of process plant models through digital watermarking[J]. Multimedia Tools and Applications, 2014, 73(3): 1687-1707
- [8] Su Z Y, Li W Q, Kong J S, *et al.* Watermarking 3D CAPD models for topology verification[J]. Computer-Aided Design, 2013, 45(7): 1042-1052
- [9] Su Z Y, Zhou L, Li W Q, *et al.* Topology authentication for CAPD models based on Laplacian coordinates[J]. Computers & Graphics, 2013, 37(4): 269-279
- [10] Huang Xiaojian. Research on fast rendering techniques in engineering CAD[D]. Beijing: Institute of Computing Technology, Chinese Academy of Sciences, 2002(in Chinese)
(黄晓剑. 工程 CAD 中的快速绘制技术研究[D]. 北京: 中国科学院计算技术研究所, 2002)
- [11] Peng F, Liu Y, Long M. Reversible watermarking for 2D CAD engineering graphics based on improved histogram shifting[J]. Computer-Aided Design, 2014, 49: 42-50
- [12] Lee S H, Kown K R. CAD drawing watermarking scheme[J]. Digital Signal Processing, 2010, 20(5): 1379-1399
- [13] Peng Fei, Huang Xiaofeng, Deng Tao, *et al.* A blind watermarking algorithm for 2D engineering graphics based on entity substitution and mean value of wavelet coefficients[J]. Journal of Computer Research and Development, 2010, 47(8): 1346-1353(in Chinese)
(彭飞, 黄小峰, 邓韬, 等. 一种基于实体替换与小波均值的二维工程图盲水印算法[J]. 计算机研究与发展, 2010, 47(8): 1346-1353)
- [14] Li Li, Li Qiang, Fang Xiaojun, *et al.* Reversible watermarking algorithm based on difference expansion for 2D engineering drawing[J]. Journal of Image and Graphics, 2010, 15(3): 372-377(in Chinese)
(李黎, 李强, 方晓君, 等. 基于差值扩大的 2 维工程 CAD 图可逆水印算法[J]. 中国图象图形学报, 2010, 15(3): 372-377)
- [15] Chen Yongfu, Cui Hongxing, Wang Yanwei, *et al.* A semi-

- fragile watermarking algorithm for copyright protection and tamper localization of engineering graphics[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2014, 26(8): 1240-1250(in Chinese)
(陈永府, 崔洪星, 王彦伟, 等. 面向工程图版权保护和修改定位的半脆弱水印算法[J]. *计算机辅助设计与图形学学报*, 2014, 26(8): 1240-1250)
- [16] Wang N N, Men C G. Reversible fragile watermarking for 2D vector map authentication with localization[J]. *Computer-Aided Design*, 2012, 44(4): 320-330
- [17] Lee S H, Hwang W J, Kwon K R. Polyline curvatures based robust vector data hashing[J]. *Multimedia Tools and Applications*, 2014, 73(3): 1913-1942
- [18] Wu H T, Huang J W, Shi Y Q. A reversible data hiding method with contrast enhancement for medical images[J]. *Journal of Visual Communication and Image Representation*, 2015, 31: 146-153
- [19] Zhao Y, Wang S Z, Zhang X P, *et al.* Robust hashing for image authentication using Zernike moments and local features[J]. *IEEE Transactions on Information Forensics and Security*, 2013, 8(1): 55-63
- [20] Pan Zhigeng, Sun Shusen, Li Li. An overview of 3D model watermarking[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2006, 18(8): 1103-1110(in Chinese)
(潘志庚, 孙树森, 李 黎. 三维模型数字水印综述[J]. *计算机辅助设计与图形学学报*, 2006, 18(8): 1103-1110)
- [21] Sakiyama A, Tanaka Y. Oversampled graph Laplacian matrix for graph filter banks[J]. *IEEE Transactions on Signal Processing*, 2014, 62(24): 6425-6437
- [22] Wang W B, Zheng G Q, Yong J H, *et al.* A numerically stable fragile watermarking scheme for authenticating 3D models[J]. *Computer-Aided Design*, 2008, 40(5): 634-645
- [23] Peng F, Guo R S, Li C T, *et al.* A semi-fragile watermarking algorithm for authenticating 2D CAD engineering graphics based on log-polar transformation[J]. *Computer-Aided Design*, 2010, 42(12): 1207-1216