

基于二维码图书版权保护的数字水印算法^①



张明, 杨辉, 黄炳家, 郑秋梅

(中国石油大学(华东), 青岛 266580)

摘要: 版权保护一直是图书出版领域的难题之一. 根据二维码和主成分分析(PCA)算法的特点, 提出一种基于二维码图书版权保护的新的数字水印算法, 算法首先对二维码载体图像进行主成分分析(PCA), 提取出既含有高频分量又含有低频分量的主成分系数, 采用 Arnold 算法将版权信息水印加密干扰后嵌入到提取出的主成分系数中. 实验结果显示, 新算法在几何攻击、噪声、图像亮度和对比度的增减等攻击都有很好的鲁棒性.

关键词: 版权保护; 二维码; PCA; Arnold 置乱; 鲁棒性

引用格式: 张明, 杨辉, 黄炳家, 郑秋梅. 基于二维码图书版权保护的数字水印算法. 计算机系统应用, 2019, 28(9): 190-195. <http://www.c-s-a.org.cn/1003-3254/7064.html>

Digital Watermarking Algorithm Based on Book Copyright Protection of QR Code

ZHANG Ming, YANG Hui, HUANG Bing-Jia, ZHENG Qiu-Mei

(China University of Petroleum, Qingdao 266580, China)

Abstract: Copyright protection has always been one of the difficult issues in the field of book publishing. Digital watermarking has been widely used as an important method of copyright protection. According to the characteristics of the QR code and PCA algorithm, this study proposes a digital watermarking algorithm based on QR code book copyright protection. The algorithm firstly uses Principle Components Analysis (PCA) to decorrelate the QR code image pixel to obtain the principle component which contains both high frequency and low frequency, then embeds the copyright information watermark after using Arnold scrambling to these principal component factor. The experimental results show that the new algorithm has strong robustness against geometric attacks, noise, image brightness and contrast increase or decrease attacks.

Key words: copyright protection; QR code; Principle Components Analysis (PCA); Arnold scrambling; robustness

随着图书电子资源的发展, 图书版权保护一直是图书出版领域的一个难题^[1], 同时, 数字水印技术作为近年来版权保护的重要技术受到越来越多的关注. 现在的数字水印算法主要基于空间域和变换域两种, 但是仍然难以实现水印嵌入容量、不可见性和鲁棒性三者的平衡^[2]. 文献^[3]提出了基于离散时间混沌系统的新的离散小波变换水印算法, 不仅计算量大, 而且对几何攻击鲁棒性效果差. 文献^[4]提出了将二维码奇异值分解

后嵌入到载体图像中, 虽然在无攻击情况下能正确识别出二维码图像, 但算法在受噪声和对比度变化攻击后鲁棒性很差. 文献^[5]提出了一种结合离散小波变换、离散余弦变换和奇异值分解的混合域水印算法, 过程繁琐, 且对几何攻击鲁棒性也表现不好. 二维码技术作为一项研究热点, 以信息存储量大, 可靠性高, 自动纠错能力强^[4]等优点已经广泛应用于各行各业中. 结合二维码的特点, 笔者提出了一种基于 Arnold 算法

① 基金项目: 国家自然科学基金(61305008)

Foundation item: National Natural Science Foundation of China (61305008)

收稿时间: 2019-03-04; 修改时间: 2019-03-29; 采用时间: 2019-04-01; csa 在线出版时间: 2019-09-05

和主成分分析 (PCA) 的二维码数字水印算法, 将图书鉴权图像信息置乱后嵌入到图书二维码图像信息中, 能有效的实现图书版权的保护。

1 相关技术

1.1 二维码原理

二维码是一种按特定规则排列的黑白相间的模块图形, 其中每个模块代表一个单元, 每个单元又被编码为 1 bit 的数据, 这些数据都是“0”, “1”二进制比特流, 能通过特定设备进行识别和读写。

二维码主要包括两部分内容, 数据区和检测区^[6] (如图 1)。数据区主要存储数据, 检测区主要用于符号定位和信息结构确认。二维码的结构不仅保证了自己具有一定的校验功能, 而且识别范围广, 处理速度快, 能够对不同行的信息进行自动识别, 对图像的几何变换能够自动处理等功能。同时二维码可以编码任何可读信息, 例如图片、声音和文本, 并以二维条形码形式显示出来, 也可表示为图像数据, 极大方便了信息的处理^[7]。结合二维码自动纠错能力强、嵌入容量大、安全性高的特点, 将二维码应用于数字水印版权保护中将起到事半功倍的效果。

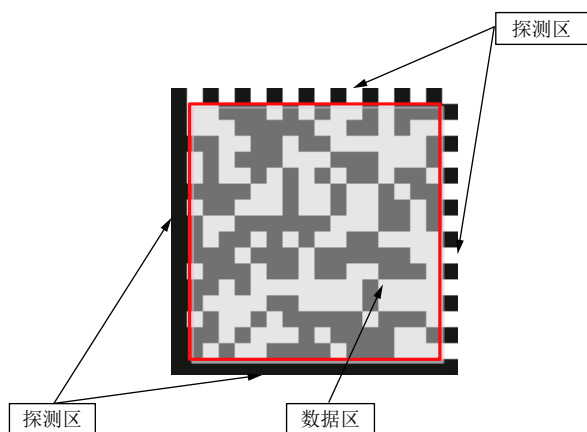


图 1 二维码探测区与数据区

1.2 Arnold 变换原理

目前图像变换主要基于像素的大小和像素的位置两种形式, 根据这种特点图像置乱加密也主要分为两种: 像素大小的加密和像素位置的加密。Arnold 变换^[8]是一种经典的位置加密方法, 图像经过某种运算形成表面杂乱无章, 无法解释但隐含一定规律的图像, 这种变换可以有效抵抗图像在几何变换方面的攻击, 同时,

Arnold 变换存在的周期性让图像恢复变得简单易行。

Arnold 变换公式如下: 设某图像某点的坐标为 (x, y) , 然后将点 (x, y) 变换到另一点 (x', y') 的变换规则为:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod (N)$$

其中, N 为正方形的边长。

则 Arnold 反变换为:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} \bmod (N)$$

数学方法证明当进行 n 次 Arnold 变换后, 即:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} x \\ y \end{bmatrix} \bmod (N)$$

Arnold 反变换为:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}^n \begin{bmatrix} x' \\ y' \end{bmatrix} \bmod (N)$$

式中, (x, y) 和 (x', y') 分别是图像置乱前后的像素位置。

与其他置乱方法相比, Arnold 变换做到了水印图像之间相关性的真正消除, 且加密方法简单易于实现, 因此 Arnold 置乱得到了普遍的应用。

1.3 主成分分析 (PCA) 的基本原理

PCA 是一种多元降维算法^[9,10], 通过一定的运算把某一些相关或相似数据变量转换成新的数据变量, 这些新变量两两不相关且在反映原始图像信息方面保持客观不变性, 新变量具有随方差逐渐递减的特点。

设有 n 个样本 X_1, \dots, X_p , p 维向量 $x = (x_1, \dots, x_p)^T$, $i = 1, 2, \dots, n, n > p$, 构造初始矩阵如下:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & \cdots & X_{pn} \end{bmatrix} \quad (1)$$

PCA^[10,11] 的计算基本过程如下:

(1) 标准化过程. 对矩阵 X 进行标准化处理:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \dots, n, j = 1, 2, \dots, p. \quad (2)$$

其中,

$$\bar{x}_j = \frac{\sum_{i=1}^n x_{ij}}{n}, s_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}$$

式中, \bar{x}_j 为均值, s_j 为 X 的分量的方差。

(2) 计算矩阵 Z 的相关系数矩阵 R .

$$R = [r_{ij}]_{p \times p} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix} \quad (3)$$

其中,

$$r_{ij} = \frac{\sum Z_{ij}Z_{ij}}{n-1}, \quad i, j = 1, 2, \dots, p.$$

(3) 样本相关矩阵 R 的特征方程. 根据下式求得 p 个特征值:

$$|R - \lambda I_p| = 0. \quad (4)$$

排列顺序为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. 根据特征值 λ_i 求出特征向量 $e_i (i = 1, 2, \dots, p)$, 然后根据 e_i 形成特征系数矩阵 $U = (e_1, e_2, \dots, e_p)^T$.

(4) 确定主成分数. 通常意义上规定, 每个主成分在所有样本分析中所占的百分数称为贡献率(CR), 相应的主成分总和对各个成分的方差之和的贡献率称为累积贡献率(ACR).

设 λ_i 表示第 i 个特征值, 则相应的第 i 个主元素的 CR(r) 为

$$CR(r) = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \quad (5)$$

其中, λ_i 为特征值, $i = 1, 2, \dots, p$.

综合前面公式求得前 m 个主成分的 ACR(m) 为:

$$ACR(m) = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (6)$$

在实际应用中, 一般采取 $ACR \geq 85\%$ 确定 m 值, 这样才能使所提取信息的达到好的利用率.

(5) 变换真正的主成分. 按下式计算主成分:

$$F_j = U_j^T Z, \quad j = 1, 2, \dots, m \quad (7)$$

为了将水印信息嵌入到二维码载体图像最有效的位置, 本文采用 PCA 从二维码载体图像中选择提取出了最有效的图像主要系数. 与其他的频域变换不同, PCA 提取的主成分系数同时包含图像的高频分量和低

频分量. 因此, 水印嵌入这些系数中可以充分的避开普通频域算法的各种难题. 在利用二维码编码特性以及 PCA 优点的基础上, 通过选取适当的水印嵌入算法, 并结合行之有效的嵌入系数和嵌入强度, 大大提高了水印算法的鲁棒性.

2 基于二维码图书版权保护的数字水印算法

2.1 算法模型

算法嵌入: 原始水印采用 Arnold Cat 变换进行初步加密置乱, 得到水印图像 W , 将原始二维码图像分块, 然后进行主成分分析, 得到有效的系数, 根据实验筛选出最佳的嵌入系数, 利用加法原则把水印图像 W 嵌入到这些系数中, 最后实现主成分分析逆变换, 获得嵌入水印的载体图像.

算法提取: 对带有水印的载体图像进行 PCA 分析, 得到新的主要成分, 根据前面嵌入算法中的参数和加法原则进行变换, 得到加密水印, 最后通过 Arnold Cat 逆变换可得到原始水印图像.

算法流程图如图 2.

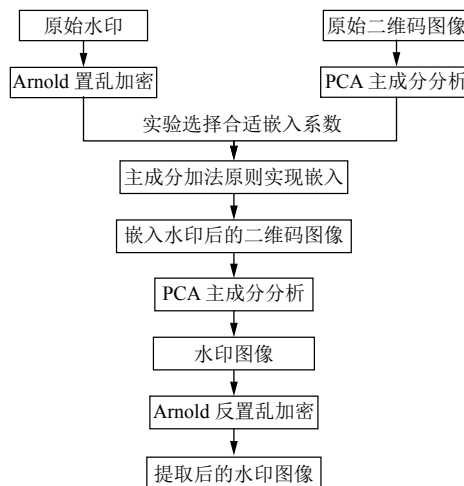


图 2 水印算法流程图

2.2 水印的嵌入过程

在嵌入原始水印前, 本文采用 Arnold Cat^[12]变换对图像进行初步加密置乱, 算法不仅计算简单易行, 而且能有效抵抗裁剪等几何攻击, 再通过其变换的周期性, 恢复原始图像. 图 3 分别显示了原始水印以及置乱后的水印, 算法的密钥 $key = 0.2345$.



图3 原始水印图像和加密置乱后水印图像

设原始二维码图像是 I , 水印是 W , 加密置乱后的水印是 W' . 具体嵌入过程如下:

(1) 首先, 把二维码图像进行 8×8 块处理, 最终分成许多子块 $I_n (n = 1, 2, \dots, 4096)$. 然后标准化每个子块 I_n 并生成矩阵 $Z(i, j)$.

(2) 根据矩阵 $Z(i, j)$ 计算出相应的相关系数矩阵 $R(i, j)$.

(3) 根据相关系数矩阵 R 的特征方程, 得到 P 个特征根值, 然后依次按照降序排列, 即 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. 在获得特定值 λ_i 后, 就可以根据 λ_i 计算求得相应的特征向量 $e_i (i = 1, 2, \dots, p)$, 然后综合利用 e 组成相应的矩阵为 $U = (e_1, e_2, \dots, e_p)^T$.

(4) 变换形成真正的主成分. 计算公式为

$$y_j = U_j^T Z, j = 1, 2, \dots, m \quad (8)$$

(5) 水印嵌入.

$$Y' = y + aw \quad (9)$$

式中, $i = 1, 2, \dots, m$, m 为提取的主成分个数; a 为水印嵌入强度; y 和 Y' 分别是原始的主要成分以及将水印嵌入后的主要成分.

根据嵌入处理过程对主成分系数 Y' 进行逆变换, 生成含有水印的载体图像 I^W .

步骤 (1) 将二维码载体图像层层分块并进行了标准化处理, 为后面进行 PCA 主成分分析奠定了基础. 步骤 (2)–(4) 是将每块的图像实施主成分分析, 分解找出每块的主成分系数, 它们是原始二维码图像低频分量和高频分量的结合体, 把水印充分合理的嵌入到这些系数中可以有效提高算法的鲁棒性. 最后一步是根据实验筛选出最佳的嵌入系数, 然后根据式 (9) 实现水印的合理嵌入.

笔者根据主成分自身拥有的特点, 选用简单的加法原理实现水印的有效嵌入. 经过实验证明, 不同的水印嵌入系数, 对水印算法的鲁棒性和不可见性有不同的效果. 当系数 a 越大, 水印鲁棒性表现会越好, 但是相

应的不可见性表现会越差. 所以系数 a 必须通过实验选取, 根据实验 a 选择 0.03.

2.3 水印的提取过程

在水印提取步骤中, 根据 Arnold Cat^[12] 加密后密钥 key , 强度系数 a 获得完整水印. 具体步骤如下:

(1) 首先将原始二维码图像 I 依次根据嵌入过程中的前 4 步骤计算得到原始主要成分 $y_i (i = 1, 2, \dots, m)$.

(2) 把包含水印的图像 I^W 同样依次根据嵌入步骤中的前 4 步骤分析获得新的主要成分 $Y'_i (i = 1, 2, \dots, m)$.

(3) 根据如下公式提取水印.

$$w = \frac{Y' - y}{a} \quad (10)$$

式中, $i = 1, 2, \dots, m$, m 是提取的主成分数量, a 是水印嵌入系数, y_i 是嵌入前载体图像的主成分系数, Y' 是水印图像的主成分系数, w 是提取后的水印.

原始二维码图像自带信息容量大, 可靠性能高, 自动纠错能力强等特点, 不仅能提高嵌入容量而且能有效抵抗一些攻击. 同时, 对二维码载体图像进行 PCA 分析, 获得图像的主要成分, 这些主要成分可以最佳的表示图像主要特征, 它们不仅代表高频分量而且含有低频分量, 将水印和主成分合理的融合, 可以最优化的实现水印不可见性和鲁棒性平衡. 结合以上两种特点, 新算法水印鲁棒性表现很强.

3 实验与结论

3.1 数字水印评价标准

3.1.1 峰值信噪比(PSNR)

峰值信噪比是评价两幅图像相似程度最通用的指标. $PSNR$ ^[11-14] 值越大, 表示两幅图像越接近. 计算公式为:

$$PSNR = 10 \times \lg \left[\frac{MN \max(I)^2}{\sum_{i=1}^M \sum_{j=1}^N (I - I')^2} \right] \quad (11)$$

式中, I 是原始载体图像, I' 是含有水印后的图像, M , N 表示图像的大小.

3.1.2 归一化相关系数(NC)

归一化相关系数是评价原始水印与提取水印之间相似度的有效指标. NC ^[11-14] 值越大越好, 计算公式为:

$$NC = \frac{\sum_{i=1}^L w(i) * w'(i)}{\sqrt{\sum_{i=1}^L w^2(i)} \sqrt{\sum_{i=1}^L w'^2(i)}} \quad (12)$$

式中, $w(i)$ 是原始水印, $w'(i)$ 是算法提取出的水印, L 代表水印的长度.

4 实验结果及分析

通过 $PSNR$ 评估新算法的有效鲁棒性, 图 4 是二维码载体图像以及含有水印的二维码图像, 通过计算求得两者之间的 $PSNR=42.8876$. 说明两图像相似程度高^[2-5], 所以新算法有效的实现了水印不可见性低的效果.

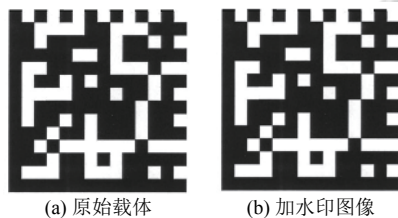


图 4 原始载体图像和加水印后图像

在本文中, 我们使用文献[3]提出的水印算法做比较, 依次进行无攻击实验, 以及例如裁剪, 旋转, 噪声和图像变化等攻击实验.

表 1 各种攻击后提取的水印的 NC 值

| 攻击种类 | 参数 | 本文算法 | 文献[3] |
|--------|------------|--------|--------|
| 无攻击 | | 0.9864 | 1.0000 |
| 裁剪 | 左上角 1/5 | 0.9688 | 0.8054 |
| 裁剪 | 左上角 1/4 | 0.9621 | 0.7633 |
| 旋转 | 30° | 0.9578 | 0.6023 |
| 旋转 | 45° | 0.9477 | 0.5954 |
| 积性噪声 | 0.01 | 0.9856 | 0.8079 |
| 高斯噪声 | 0.01 | 0.9825 | 0.6709 |
| 图像增亮 | [0.4, 1] | 0.9773 | 0.6978 |
| 图像变暗 | [0, 0.95] | 0.9864 | 0.8488 |
| 直方图均衡化 | | 0.9711 | 0.8812 |
| 增加对比度 | [0.2, 0.6] | 0.9560 | 0.5991 |
| 降低对比度 | [0.2, 0.8] | 0.9788 | 0.6984 |

图 5 显示了原始水印图像以及提取的水印图像, 前者经过 PCA 获得的系数是表示图像主要特征的对角矩阵, 把鉴权水印信息和获得的系数合理的融合能充分的抵抗攻击, 但是相应的, 算法进行变换以及逆变换的过程中, 某些信息会有所丢失. 图 6 至图 10 是本文的算法和文献[3]算法在裁剪, 旋转, 高斯噪声, 图像

变化和对比度变化等攻击后提取的水印图像. 可以看出只有在无攻击条件下, 新水印算法的性能比文献[3]水印算法的性能稍差, 但是在其他攻击例如噪声、对比度变化等攻击后, 新算法都优于文献[3], 而且在几何攻击旋转以及裁剪的 NC 值明显优于文献[3]. 本文通过结合二维码的特征和 PCA 算法, 将水印合理有效的嵌入到最优主成分中, 最大限度的实现了保持水印结构完整性的可能.

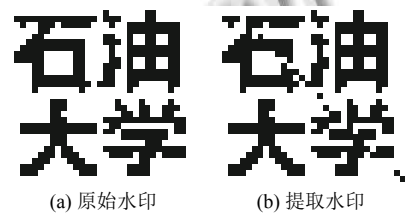


图 5 原始水印图像和提取的水印图像

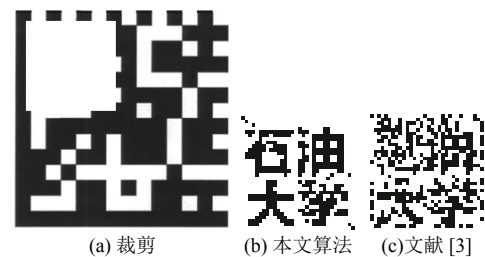


图 6 裁剪攻击

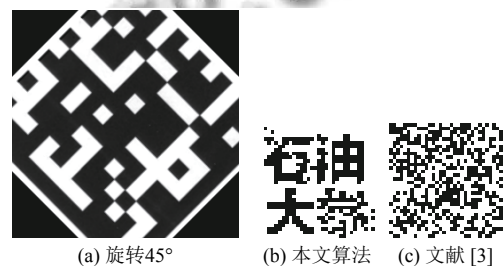


图 7 旋转攻击

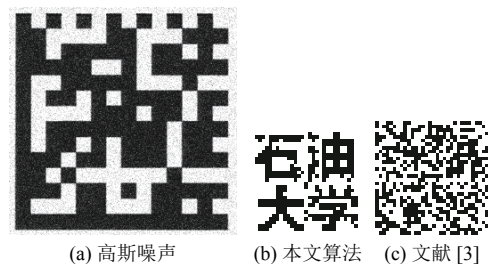


图 8 高斯噪声攻击

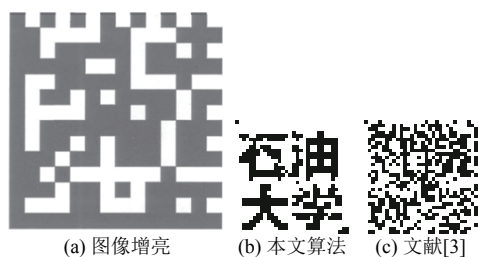


图9 图像增亮攻击

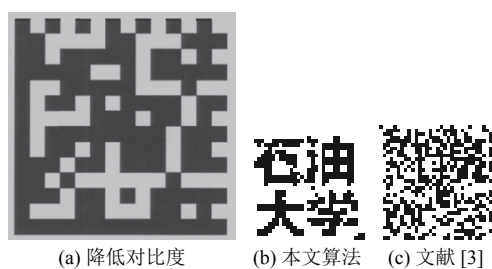


图10 降低对比度攻击

5 结束语

由于二维码信息存储量大,可靠性能高,自动纠错能力强等特点,本文结合PCA算法提出了一种基于二维码版权保护的数字水印算法,将Arnold加密后的鉴权信息合理嵌入到二维码载体中.实验结果表明,与文献[3]提出的水印算法相比,新算法不仅在旋转、裁剪等几何攻击方面表现出了很强的鲁棒性,在图像亮度增减、对比度的增减和噪声的增加等攻击方面,该算法也大大提高了水印的鲁棒性.

参考文献

- 1 兰丽专. 图书出版版权保护的困境与出路. 传播与版权, 2015, (2): 178-179.
- 2 郑秋梅, 顾国民, 王玉菲, 等. 一种新的抗几何攻击的数字

- 3 水印算法. 中国石油大学学报(自然科学版), 2012, 36(1): 188-192. [doi: 10.3969/j.issn.1673-5005.2012.01.034]
- 3 Hannoun K, Hamiche H, Lahdir M, *et al.* A novel DWT domain watermarking scheme based on a discrete-time chaotic system. IFAC-Papers On Line, 2018, 51(33): 50-55. [doi: 10.1016/j.ifacol.2018.12.089]
- 4 葛娅敬, 赵礼峰. 基于奇异值分解的二维码加密算法. 计算机科学, 2018, 45(11A): 342-343, 346. [doi: 10.11896/j.issn.1002-137X.2018.11A.069]
- 5 孙汉卿, 李喜艳, 王桂芝, 等. DWT-DCT-SVD 域的水印置乱新研究. 激光杂志, 2019, 40(2): 110-113.
- 6 郑秋梅, 金萧, 顾国民, 等. 一种基于 Data Matrix 的数字水印算法. 中国石油大学学报(自然科学版), 2015, 39(1): 188-193. [doi: 10.3969/j.issn.1673-5005.2015.01.028]
- 7 龙陈锋, 雷坚, 杨鑫, 等. 基于数字水印—二维码标签的黄茶追溯系统研究. 湖南农业大学学报(自然科学版), 2015, 41(5): 565-568.
- 8 徐潇, 马峻, 莫凡珣, 等. 基于计算全息和 Arnold-混沌技术的三维信息分级加密. 激光杂志, 2018, 39(11): 66-70.
- 9 周非, 夏鹏程. 基于主成分分析和卡方距离的信号强度差指纹定位算法. 计算机应用, 2019, 39(5): 1405-1410.
- 10 张明. 基于 DWT 变换和 PCA 主成分分析的数字水印算法研究[硕士学位论文]. 青岛: 中国石油大学(华东), 2014.
- 11 郑秋梅, 张明, 王风华, 等. 基于 PCA 和 DWT 的强鲁棒数字水印算法. 中国石油大学学报(自然科学版), 2016, 40(1): 177-182. [doi: 10.3969/j.issn.1673-5005.2016.01.025]
- 12 Wang JZ, Wang YL, Wang MQ. Periodicity and Application for a kind of n-dimensional Arnold-type Transformation. Proceedings of 2007 IEEE Intelligence and Security Informatics. New Brunswick, NJ, USA. 2007. 375.
- 13 徐江峰, 张守强. 基于 QR 码的 DWT-DCT 数字水印算法. 计算机应用研究, 2018, 35(5): 1540-1544. [doi: 10.3969/j.issn.1001-3695.2018.05.056]
- 14 张劲松, 杨玫, 周立新. 基于 DWT 的图像内容半脆弱水印认证算法. 科技通报, 2017, 33(6): 192-195.