

可有效抵抗一般性几何攻击的数字水印检测方法

徐紫涵¹ 王向阳¹

摘要 以回归型支持向量机理论为基础, 结合性能稳定的伪 Zernike 矩和 Krawtchouk 矩, 提出了一种可有效抵抗一般性几何攻击的强鲁棒数字图像水印检测算法. 该算法首先选取图像的低阶 Krawtchouk 矩作为特征向量, 然后利用 SVR 对几何变换参数进行训练学习并对待检测图像进行数据预测, 最后对其进行几何校正并提取水印信息. 仿真实验结果表明, 该数字图像水印检测算法不仅具有较好的不可感知性, 而且对常规信号处理和一般性几何攻击均具有较好的鲁棒性.

关键词 图像水印, 几何攻击, 支持向量机, 伪 Zernike 矩, Krawtchouk 矩

中图分类号 TP391

An SVR Based Image Watermarking Detection Algorithm against Geometric Attacks

XU Zi-Han¹ WANG Xiang-Yang¹

Abstract According to the support vector regression (SVR), a new image watermarking detection algorithm against geometric attacks is proposed in this paper, in which the steady pseudo-Zernike moments and Krawtchouk moments are utilized. The main steps of watermark detecting procedure include: 1) some low-order Krawtchouk moments are calculated as the eigenvectors; 2) the appropriate kernel function is selected for training, and an SVR training model can be obtained; 3) the actual output is predicted by using the well trained SVR; 4) the digital watermark is extracted from the corrected test image. Experimental results show that the proposed watermarking detection algorithm is not only robust against common signals processing, but also robust against some geometric attacks.

Key words Image watermarking, geometric attack, support vector regression (SVR), pseudo-Zernike moment, Krawtchouk moment

数字水印 (Digital watermarking) 技术发展至今, 绝大多数图像水印算法能够对抗常规的信号处理和简单的全局仿射变换, 但无法有效抵抗一般性几何攻击, 如剪切、行列去除、局部弯曲、几何变换组合等. 因此, 抗一般性几何攻击的高度鲁棒数字图像水印算法研究仍然是一项富有挑战性的工作.

人们主要采用三种措施设计抗一般性几何攻击图像水印方案, 分别为构造几何不变量, 隐藏模板

或可识别结构, 利用原始数据重要特征^[1-4]. Dong 等^[2-4] 分别采用图像归一化、Fourier-Mellin 变换、广义 Radon 变换、Zernike 矩等构造图像几何不变量. 文献 [5-7] 相继提出了基于图像特征的数字水印方案. Fu^[8] 则利用支持向量机 (Support vector machine, SVM) 做了大量的工作. 先进的 SVM 理论为几何攻击问题提供了可能的解决方向, 但现有 SVM 图像水印方案普遍存在抗攻击能力差, 透明性低和鲁棒性弱等问题.

本文以回归型支持向量机 (Support vector regression, SVR) 理论为基础, 结合性能稳定的伪 Zernike 矩和 Krawtchouk 矩相关知识, 提出了一种可有效抵抗一般性几何攻击的强鲁棒数字图像水印方案. 仿真实验表明, 该方案不仅具有较好的不可感知性, 而且对常规信号处理及几何攻击均具有较好的鲁棒性.

1 数字水印的嵌入

本文首先将原始载体映射变换到极坐标下, 然后结合伪 Zernike 矩的旋转不变特性, 计算出伪 Zernike 矩, 最后选取部分低阶伪 Zernike 矩, 并采用量化调制策略将水印信息嵌入到伪 Zernike 矩幅值中.

收稿日期 2007-09-18 收修改稿日期 2008-01-12
Received September 18, 2007; in revised form January 12, 2008
国家自然科学基金 (60773031, 60873222), 计算机软件新技术国家重点实验室 (南京大学) 开放基金 (A200702), 信息安全国家重点实验室 (中国科学院软件研究所) 开放基金 (03-06), 图像处理与图像通信江苏省重点实验室 (南京邮电大学) 开放基金 (ZK205014), 江苏省计算机信息处理技术重点实验室 (苏州大学) 开放课题基金 (KJS0602) 资助
Supported by National Natural Science Foundation of China (60773031, 60873222), the Open Foundation of State Key Laboratory for Novel Software Technology (Nanjing University) (A200702), the Open Foundation of State Key Laboratory of Information Security of China (03-06), the Open Foundation of Key Laboratory of Image Processing and Image Communication (Nanjing University of Posts and Communications) (ZK205014), the Open Foundation of Jiangsu Province Key Laboratory for Computer Information Processing Technology (Soochow University) (KJS0602)

1. 辽宁师范大学计算机与信息技术学院, 大连 116029
1. School of Computer and Information Technology, Liaoning Normal University, Dalian 116029
DOI: 10.3724/SP.J.1004.2009.00023

设原始载体为 256 级灰度图像 $F = \{f(x, y), 1 \leq x \leq M, 1 \leq y \leq N\}$, 其中, $f(x, y)$ 表示原始载体第 x 行、第 y 列的像素灰度值, 则整个数字水印嵌入过程的关键步骤描述如下:

步骤 1. 水印产生. 由密钥 Key 1 产生一个伪随机序列 $W = \{w_i, i = 1, \dots, L\}$ 作为数字水印信息, 其中, L 为数字水印的大小, $w_i \in \{0, 1\}$.

步骤 2. 计算原始图像的伪 Zernike 矩. 伪 Zernike 矩是一组正交复数矩, 它是通过将图像映射到一组基函数上得到的, 这个基函数称为伪 Zernike 矩的基, 记为 $\{V_{mn}(x, y)\}$. 这组基构成了单位圆 ($x^2 + y^2 \leq 1$) 上的一组完备正交集, 定义为

$$V_{mn}(x, y) = V_{mn}(r, \theta) = R_{mn}(r) \exp(jn\theta) \quad (1)$$

式中 m 为非负整数, n 为整数, 两者满足 $|n| \leq m$. r 、 θ 分别为极坐标下图像像素的半径和角度, $r = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(y/x)$. $R_{mn}(r)$ 为径向多项式 (Radial polynomial), 其定义如下

$$R_{mn}(r) = \sum_{s=0}^{m-|n|} \frac{(-1)^s (2m+1-s) r^{m-s}}{s!(m+|n|+1-s)!(m-|n|-s)!} \quad (2)$$

阶数为 m , 重复度为 n 的伪 Zernike 矩定义为

$$Z_{mn} = \frac{m+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) [V_{mn}(x, y)]^* dx dy \quad (3)$$

式中, $m = 0, 1, 2, \dots$; $f(x, y)$ 是图像亮度函数; “*” 表示复数共轭; n 是整数并且满足 $0 \leq |n| \leq m$.

步骤 3. 选择伪 Zernike 矩. a) 选择阶数较低的伪 Zernike 矩. 因为当阶数高于某一数值 $M_{\max}$ 时, 伪 Zernike 矩计算将不再准确. 本文选取 $M_{\max} = 20$ (经过实验获得的最大迭代次数, 超过此数值效果不好); b) 重复度为 $n = 4i$ ($i = 0, 1, 2, \dots$) 的伪 Zernike 矩存在微小计算误差, 故不适合嵌入水印. 显然, 可用于数字水印嵌入的伪 Zernike 矩集合为

$$S = \{Z_{mn}, m \leq M_{\max}, n \geq 0, n \neq 4i\} \quad (4)$$

为了提高系统安全性能, 我们利用密钥 Key 2 从伪 Zernike 矩集合 S 中随机选择 L 个伪 Zernike 矩 $Z = (Z_{p_1 q_1}, \dots, Z_{p_L q_L})$ 用于水印嵌入. 设其对应的伪 Zernike 矩幅值为 $A = (A_{p_1 q_1}, \dots, A_{p_L q_L})$.

步骤 4. 数字水印嵌入. 本文采用量化调制伪 Zernike 矩幅值的方法实现水印信号嵌入, 量化规则为

$$A'_{p_i q_i} = \left\lceil \frac{A_{p_i q_i} - d(w_i)}{\Delta} \right\rceil \Delta + d(w_i), i = 1, 2, \dots, L \quad (5)$$

这里, $\lceil \cdot \rceil$ 是四舍五入操作, Δ 是量化步长, $d(\cdot)$ 是通过密钥 Key 3 产生的量化函数, 且满足 $d(1) = \Delta/2 + d(0)$, $d(0) \in [0, 1]$.

需要说明的是, 量化伪 Zernike 矩幅值 $A = (A_{p_1 q_1}, \dots, A_{p_L q_L})$ 时, 如果 $q_i \neq 0$, 应同时量化它的共轭矩阵幅值 $A_{p_i, -q_i}$, 以保证其具有相同幅值.

步骤 5. 含水印图像生成. 含水印图像由两部分合并而成:

第一部分是未被修改的伪 Zernike 矩所重构的数字图像 $f^*(x, y)$

$$f^*(x, y) = f_O(x, y) - f_Z(x, y) \quad (6)$$

这里, $f_O(x, y)$ 是原始载体图像, $f_Z(x, y)$ 是待修改伪 Zernike 矩的重构部分图像, 且有

$$f_Z(\cdot) = \sum_{i=1}^L Z_{p_i q_i} V_{p_i q_i}(\cdot) + Z_{p_i, -q_i} V_{p_i, -q_i}(\cdot) \quad (7)$$

第二部分是已修改伪 Zernike 矩的重构部分图像 $f_{Z'}(x, y)$

$$f_{Z'}(\cdot) = \sum_{i=1}^L Z'_{p_i q_i} V_{p_i q_i}(\cdot) + Z'_{p_i, -q_i} V_{p_i, -q_i}(\cdot) \quad (8)$$

将上面两部分合并, 即得到含水印图像

$$f'(x, y) = f^*(x, y) + f_{Z'}(x, y) \quad (9)$$

2 数字水印的检测

本文结合 Krawtchouk 矩相关知识, 给出了一种基于 SVR 几何校正的数字水印检测算法. 该算法首先选取图像的低阶 Krawtchouk 矩作为特征向量, 并通过 SVR 对旋转、缩放、平移等几何变换参数进行训练学习, 以获得 SVR 训练模型. 然后利用 SVR 训练模型对待检测图像进行数据预测, 并结合预测输出结果对其进行几何校正. 最后从已校正数字图像内提取出水印信息.

2.1 特征向量的构造

由数字图像相关理论知, 低阶图像矩可以很好地反映整幅图像模式. 考虑到本文重点讨论旋转、缩放、平移等整体变换形式, 故以下采用 Krawtchouk 矩的低阶矩 4 个特征 $Q_{00}, Q_{01}, Q_{10}, Q_{11}$ (分别记为 $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$) 来反映图像全局信息, 并进一步将其作为 SVR 训练特征向量.

设 $f(x, y)$ 表示大小为 $M \times N$ 的数字图像, 则其 $(m+n)$ 阶 Krawtchouk 矩 Q_{nm} 的定义为

$$Q_{mn} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \bar{K}_n(x; p_1, N-1) \bar{K}_m(y; p_2, M-1) f(x, y) \quad (10)$$

其中

$$\bar{K}_n(x; p, N-1) = K_n(x; p, N-1) \sqrt{\frac{\omega(x; p, N-1)}{\rho(n; p, N-1)}} \quad (11)$$

$$\omega(x; p, N) = \binom{N}{x} p^x (1-p)^{N-x} \quad (12)$$

$$\rho(n; p, N) = (-1)^n \left(\frac{1-p}{p} \right)^n \frac{n!}{(-N)_n} \quad (13)$$

$$K_n(x; p, N) = \sum_{k=0}^n a_{k,n,p} x^k = {}_2F_1 \left(-n, -x; -N; \frac{1}{p} \right) \quad (14)$$

$x, n = 0, 1, 2, \dots, N, N > 0, p \in (0, 1)$

其中, 超几何分布函数 ${}_2F_1$ 定义为

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k z^k}{(c)_k k!} \quad (15)$$

a_k 的表达式为

$$a_k = a(a+1)(a+2) \times \dots \times (a+k-1) \quad (16)$$

2.2 SVR 训练模型的获得

相对于神经网络而言, SVR 具有更强的泛化能力和学习能力. SVR 是 SVM 在回归学习中的应用, 其基本思想是: 对于给定训练样本点 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset X \times \mathbf{R}$ (其中, X 表示输入样本空间), 通过 SVR 训练回归一个函数 $f(x)$, 使由该函数求出的每个输入样本的输出值和输入样本所对应的目标值相差不超过误差 ε , 同时使回归出的函数尽量平滑.

为了获得 SVR 训练模型, 我们首先在一定范围内随机平移 (包括 X 方向、 Y 方向)、旋转和缩放原始图像 F 以产生 K 个训练样本图像 F^k , 然后计算出每个训练样本图像 F^k 的 4 个低阶 Krawtchouk 矩

$$(\mathbf{f}_1^k, \mathbf{f}_2^k, \mathbf{f}_3^k, \mathbf{f}_4^k), \quad k = 1, 2, \dots, K$$

并将其作为训练特征向量. 同时, 将相应的变换 (X 方向平移、 Y 方向平移、旋转和缩放等) 参数 $t_x^k, t_y^k, s^k, \theta^k$ ($k = 1, 2, \dots, K$) 作为训练目标值, 于是可得到训练样本

$$\Omega_k = (\mathbf{f}_1^k, \mathbf{f}_2^k, \mathbf{f}_3^k, \mathbf{f}_4^k, t_x^k, t_y^k, s^k, \theta^k), \quad k = 1, 2, \dots, K$$

考虑到平移、旋转和缩放构成对图像的线性变换, 任何一个变换对其他参数没有影响, 因此, 4 个输出之间没有耦合. 为此, 本文采用 4 个 SVR 并行结构构

成 MIMO 系统, SVR 的结构为 4 个输入, 核函数采用 RBF 径向基函数. 通过训练学习, 即可获得 SVR 训练模型.

2.3 待检测图像的几何校正

基于 SVR 的待检测图像校正过程为:

步骤 1. 计算出待检测图像 $\tilde{F}$ 的 4 个低阶 Krawtchouk 矩 $(\tilde{\mathbf{f}}_1, \tilde{\mathbf{f}}_2, \tilde{\mathbf{f}}_3, \tilde{\mathbf{f}}_4)$, 并将其作为训练特征向量.

步骤 2. 以 4 个低阶 Krawtchouk 矩 $(\tilde{\mathbf{f}}_1, \tilde{\mathbf{f}}_2, \tilde{\mathbf{f}}_3, \tilde{\mathbf{f}}_4)$ 作为输入向量, 利用已经获得的 SVR 训练模型对输出向量进行数据预测, 从而得到相应的输出向量值 (即几何变换参数 $\tilde{t}_x, \tilde{t}_y, \tilde{s}, \tilde{\theta}$).

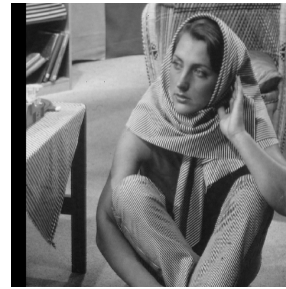
步骤 3. 利用所得到的几何变换参数 $\tilde{t}_x, \tilde{t}_y, \tilde{s}, \tilde{\theta}$ 对待检测图像 $\tilde{F}$ 进行几何校正, 从而得到待检测图像 $\tilde{F}$ 的校正结果 F^* . 图 1 给出了几何校正的结果示意图.



(a) 待检测图像 1
(a) Test image Lena



(b) 校正后的图像 1
(b) The corrected image Lena



(c) 待检测图像 2
(c) Test image Barbara



(d) 校正后的图像 2
(d) The corrected image Barbara

(c) Test image Barbara (d) The corrected image Barbara

图 1 SVR 几何校正结果示意图

Fig. 1 SVR geometric corrections

2.4 数字水印的提取

水印提取是水印嵌入的逆过程. 本文讨论的数字水印检测算法属于目标检测算法, 即在检测数字水印时不需要原始载体图像. 设校正后的待检测图像为 F^* , 则数字水印检测过程如下:

步骤 1. 计算出校正后待检测图像 F^* 的伪 Zernike 矩, 见式 (3).

步骤 2. 利用密钥 Key 2 选择 L 个伪 Zernike 矩 $Z^* = (Z_{p_1q_1}^*, \dots, Z_{p_Lq_L}^*)$ 用于水印提取. 设其对应的伪 Zernike 矩幅值为 $A' = (A'_{p_1q_1}, \dots, A'_{p_Lq_L})$.

步骤 3. 数字水印提取. 首先, 利用密钥 Key3 产生量化函数 $d(\cdot)$, 使之满足 $d(1) = \Delta/2 + d(0)$, 且 $d(0) \in [0, 1]$.

然后, 采用与嵌入过程相同的量化公式, 用两个量化函数 $d(0), d(1)$ 分别量化 $A_{p_iq_i}^*, i = 1, 2, \dots, L$

$$(A_{p_iq_i}^*)_{Q_j} = \left\lceil \frac{A_{p_iq_i}^* - d(j)}{\Delta} \right\rceil \Delta + d(j), \quad j = 0, 1 \quad (17)$$

通过式 (17), 可以得到两组向量式 $(A_{p_iq_i}^*)_0$ 和 $(A_{p_iq_i}^*)_1, i = 1, \dots, L$.

最后, 通过比较 $A_{p_iq_i}^*$ 及其两组量化式距离提取水印信息, 即

$$w_i^* = \arg \min_{j \in \{0, 1\}} ((A_{p_iq_i}^*)_{Q_j} - (A_{p_iq_i}^*))^2, \quad i = 1, \dots, L \quad (18)$$

其中, 式 (18) 可分解为以下几个具体步骤:

1) $A_{p_iq_i}^*$ 及其两组量化式距离分别定义为 $dis0 = ((A_{p_iq_i}^*)_0 - A_{p_iq_i}^*)^2$ 和 $dis1 = ((A_{p_iq_i}^*)_1 - A_{p_iq_i}^*)^2$;

2) 计算 1) 中两个距离的差值, 记为 $t = dis0 - dis1$;

3) 根据 2) 的计算结果提取水印信息, 即如果 $t < 0$, 则有 $w_i^* = 0$; 否则, 有 $w_i^* = 1$.

3 仿真试验

为了验证本文数字图像水印算法的有效性, 以下分别给出了不可感知性测试和抗攻击能力测试 (包括常规信号处理和一般性几何攻击) 的实验结果. 实验中, 所选用的原始载体分别为 $128 \times 128 \times 8$ bit 标准灰度图像 Lena、Mandrill 和 Barbara, 数字水印采用了 64 bit 的随机序列. 另外, 量化步长 $\Delta = 2.0$, 训练样本数目为 $K = 50$. SVR 训练时选用了 RBF 核函数. 图 2 (见下页) 为 Lena、Mandrill 和 Barbara 的含水印图像.

为了检测本文算法的鲁棒性能, 仿真实验分别对本文算法的含水印图像进行了一系列攻击. 表 1 和表 2 分别给出了本文算法的抗攻击能力结果 (失真率 BER).

4 结论

本文以回归型支持向量机理论为基础, 结合性能稳定的伪 Zernike 矩和 Krawtchouk 矩相关知识,

提出了一种可有效抵抗一般性几何攻击的强鲁棒数字图像水印检测方案. 其主要包括如下特点: 1) 选用抗噪能力较强的伪 Zernike 矩嵌入水印, 选用性能稳定的 Krawtchouk 矩作为训练特征, 不仅改善了 SVR 模型的工作性能, 而且提高了图像水印系统的鲁棒性; 2) 所提出的 SVR 几何校正方法可与任何其他水印嵌入方法结合, 通用性较强; 3) 该算法可有效抵抗旋转、缩放、平移、剪切、局部变形、联合攻击等多种形式的一般性几何攻击.

表 1 数字水印对常规信号处理的抵抗能力 (失真率 BER)

Table 1 The watermark detection results for common signal processing

攻击方式	Lena	Mandrill	Barbara
JPEG 压缩			
70	0	0	0
50	0	0	0
30	0	0.063	0
高斯低通滤波	0	0	0
均值滤波	0.063	0	0.031
高斯噪声	0	0.031	0
椒盐噪声	0	0	0
椒盐噪声 + 70 因子 JPEG 压缩	0	0	0
高斯滤波 + 70 因子 JPEG 压缩	0	0	0
高斯噪声 + 均值滤波	0.031	0	0.063
均值滤波 + 锐化	0	0	0.031
高斯噪声 + 高斯滤波 + 锐化	0.156	0.188	0.281

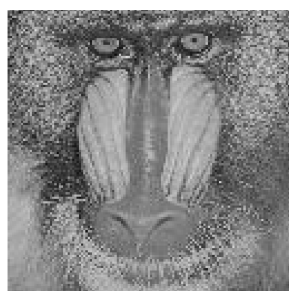
表 2 数字水印对几何攻击及联合攻击的抵抗能力 (失真率 BER)

Table 2 The watermark detection results for various attacks

攻击方式	Lena	Mandrill	Barbara
旋转			
5°	0.188	0.094	0.219
45°	0.156	0.094	0.219
缩放			
0.5	0	0	0.063
2	0	0	0
平移			
水平 5	0	0	0
水平 30	0	0	0
垂直 5	0	0	0
垂直 30	0	0	0c
横向翻转	0	0	0
局部挤压扭曲	0.031	0	0
缩放 2 + 旋转 5°	0	0	0
旋转 45° + 平移 (垂直 30)	0	0	0
JPEG 压缩 70 + 高斯噪声 + 低通滤波 + 缩放 0.5	0.031	0.063	0.094



(a) 含水印图像 Lena
(PSNR = 41.365 dB)



(b) 含水印图像 Mandrill
(PSNR = 42.106 dB)



(c) 含水印图像 Barbara
(PSNR = 40.406 dB)

(a) The watermarked image Lena

(b) The watermarked image Mandrill

(c) The watermarked image Barbara

图 2 数字水印的嵌入效果

Fig. 2 The watermarked images

References

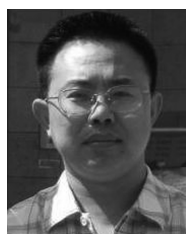
- 1 Licks V, Jordan R. Geometric attacks on image watermarking system. *IEEE Multimedia*, 2005, **12**(3): 68–78
- 2 Dong P, Brankov J G, Galatsanos N P, Yang Y Y, Davoine F. Digital watermarking robust to geometric distortions. *IEEE Transactions on Image Processing*, 2005, **14**(12): 2140–2150
- 3 Simitopoulos D, Koutsonanos D E, Strintzis M G. Robust image watermarking based on generalized radon transformations. *IEEE Transactions on Circuits and Systems for Video Technology*, 2003, **13**(8): 732–745
- 4 Parameswaran L, Anbumani K. A robust image watermarking scheme using image moment normalization. *Proceedings of World Academy of Science, Engineering and Technology*, 2006, **13**: 239–243
- 5 Wang Xiang-Yang, Wu Jun, Hou Li-Min. A feature-based digital image watermarking algorithm. *Acta Electronica Sinica*, 2007, **35**(7): 1318–1322
(王向阳, 邬俊, 侯丽敏. 一种基于图像特征点的数字水印嵌入方法. *电子学报*, 2007, **35**(7): 1318–1322)
- 6 Lee H Y, Kim H, Lee H K. Robust image watermarking using local invariant features. *Optical Engineering*, 2006, **45**(3): 1–11
- 7 Seo J S, Yoo C D. Image watermarking based on invariant regions of scale-space representation. *IEEE Transactions on Signal Processing*, 2006, **54**(4): 1537–1549
- 8 Fu Y G, Shen R, Lu H. Watermarking scheme based on support vector machine for color images. *Electronics Letters*, 2004, **40**(16): 986–987



徐紫涵 辽宁师范大学硕士研究生. 主要研究方向为数字水印技术.

E-mail: xzh0122@163.com

(XU Zi-Han Master student at Liaoning Normal University. Her main research interest is digital watermarking.)



王向阳 辽宁师范大学计算机与信息技术学院教授. 主要研究方向为信息隐藏与数字水印、多媒体信息处理技术. 本文通信作者. E-mail: wxy37@126.com

(WANG Xiang-Yang Professor at School of Computer Science and Information Technology, Liaoning Normal University. His research interest covers

information hiding, digital watermarking, and multimedia information processing. Corresponding author of this paper.)