一种基于 SVR 几何校正的数字水印检测算法

王向阳[1,2] 徐紫涵[1]

[1](辽宁师范大学计算机与信息技术学院,大连 116029)

[2](中国科学院软件研究所 信息安全国家重点实验室,北京 100039)

摘 要 以回归型支持向量机(SVR)理论基础,提出了一种可有效抵抗几何攻击的图像水印检测新算法. 该算法首先选取图像的组合矩作为特征向量,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. 仿真实验结果表明,本文算法对常规信号处理(滤波、叠加噪声、JPEG 压缩等)和几何攻击(旋转、缩放、平移、剪切等)均具有较好的鲁棒性

关键词 数字水印 几何攻击 回归型支持向量机 几何校正

中图法分类号:TP391 **文献标识码**:A **文章编号**:1006-8961(2009)06-1131-05

Image Watermarking Detection Based on SVR Geometric Correction

WANG Xiang-yang[1,2], XU Zi-han[1]

[1](*School of Computer and Information Technology, Liaoning Normal University, Dalian* 116029)

[2](*State Key Laboratory of Information Security, Institute of Software of Chinese Academy of Sciences, Beijing* 100039)

Abstract In this paper, a robust image watermarking detection based on support vector regression (SVR) is proposed. Firstly, six combined low order image moments are taken as the feature vector and the geometric transformation parameters are regarded as the training objective, the appropriate kernel function is selected for the training, and a SVR training model can be obtained. Secondly, the combined moments for test image are selected as input vector, the actual output is predicted by using the well trained SVR, and the geometric correction is performed on the test image by using the obtained geometric transformation parameters. Finally, 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 such as filtering, sharpening, noise adding, JPEG compression etc, but also robust against the geometric attacks such as rotation, translation, scaling, cropping, combination attacks, etc.

Keywords digital watermarking, geometric attacks, support vector regression, geometric correction

1 引 言

数字图像作品的知识产权保护成为迫切需要解决的关键问题。在这一背景下,数字图像水印技术日益受到关注,并已成为信息安全研究领域的一个热点[1-2]。近年来,图像水印技术研究取得了很大进展,并陆续提出了诸如空域、变换域、压缩域、基于

基金项目:国家自然科学基金项目(60773031,60873222);计算机软件新技术国家重点实验室(南京大学)开放基金项目(A200702);信息安全国家重点实验室(中国科学院软件研究所)开放基金项目(03-06);大连市科技基金项目(2006J23JH020);江苏省计算机信息处理技术重点实验室(苏州大学)开放课题基金项目(KJS0602);图像处理与图像通信"江苏省重点实验室(南京邮电大学)开放基金项目(ZK205014);辽宁省教育厅高等学校科研项目(2008351)

收稿日期:2007-07-15;**改回日期**:2007-11-01

第一作者简介:王向阳(1965~),男。教授,1995 年于吉林大学获硕士学位。主要研究方向为多媒体信息处理技术,网络信息安全技术等。E-mail:wxy37@126.com

统计学、基于生理模型等多种数字水印算法。但是，现有绝大多数图像水印方案无法有效抵抗诸如旋转、缩放、平移、剪切等几何攻击。因此，抗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[2-3]。

一般说来，对于给定的数字水印算法，水印检测器必需知道数字水印嵌入的确切位置。所谓几何攻击，并非指该种攻击能够从含水印对象中去除水印信息，而是指其能够破坏数字水印分量的同步（即改变水印嵌入位置），从而导致检测器找不到有效水印。截止到目前，人们主要采用三种措施设计抗几何攻击的图像水印方案，分别为构造几何不变量、隐藏模版、利用原始图像重要特征[3]。Dong 等人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[4-7]，并将水印信号嵌入到几何不变量内，实现了水印系统对几何攻击的鲁棒性，但该类方案抵抗混合型几何攻击（如平移 + 缩放等）能力一般，且普遍具有不可感知性较差等弱点。文献[8]~[9]通过在图像 DFT（傅里叶变换）中频区域嵌入模板信息的方式来估计并校正图像所经历的几何变换，从而实现水印检测的重同步，但现有基于模版的图像水印抵抗常规信号处理能力较差，同时水印容量也受到限制。文献[10]~[11]相继提出了基于图像特征的数字水印方案，其基本思想为：利用图像中相对稳定的特征点标识水印嵌入位置，并在与每个特征点相对应的局部区域内独立地嵌入数字水印，同时利用特征点来定位和检测数字水印，从而有效抵抗几何攻击，然而目前该类方法普遍存在特征点稳定性差且分布极不均匀等问题，严重影响了数字水印对常规信号处理的抵抗能力，同时水印容量十分有限（仅为 16bits）。

本文结合回归型支持向量机（SVR）理论，提出了一种基于 SVR 几何校正的数字水印检测算法。实验结果表明，该数字图像水印算法不仅具有良好的不可感知性，而且具有较强的抗攻击能力。

2 数字水印的嵌入

离散傅里叶变换（DFT）是一种经典而有效的信号分析工具。本文将结合离散傅里叶变换的周期性、共轭对称等诸多优良性质，给出一种 DFT 域数字水印嵌入算法。

假设原始载体为 256 级灰度图像 $\boldsymbol{F}=\{f(i,j),0\leqslant i<M,0\leqslant j<N\}$，其中 $f(i,j)$ 表示原始载体图像 $\boldsymbol{F}$ 第 i 行、第 j 列的像素灰度值。数字水印为二值图像 $\boldsymbol{W}=\{w(i,j),0\leqslant i<P,0\leqslant j<Q\}$，其中 $w(i,j)\in\{0,1\}$ 表示水印图像的第 i 行、第 j 列像素灰度值。则数字水印的嵌入过程（关键步骤）为

（1）数字水印的加密处理　首先将二值水印图像利用行扫描形成 1 维向量，并依次标号为 1 到 $P\times Q$，即得到由原二值水印图像 $\boldsymbol{W}$ 转换而来的 1 维数字水印序列 $\boldsymbol{V}$。然后利用 logistic 映射产生混沌密钥，对 1 维数字水印序列 $\boldsymbol{V}$ 进行加密处理[1]，以得到安全水印序列 $\boldsymbol{VM}$

$$\boldsymbol{VM}=\{vm(k),0\leqslant k<P\times Q,vm(k)\in\{0,1\}\} \tag{1}$$

（2）原始载体 DFT 及嵌入位置确定首先对原始载体图像实施 2 维 DFT 变换（变换原点为图像中心），得到中心化频谱，即

$$F(u,v)=DFT(\boldsymbol{F}),\ 0\leqslant u<M,0\leqslant v<N \tag{2}$$

式中，$F(u,v)$ 为复数。设 $M(u,v)=|F(u,v)|$，$F(u,v)=M(u,v)e^{i\Phi(u,v)}$，则 $M(u,v)$ 称为幅值谱，$\Phi(u,v)$ 称为相位谱。

然后在中心化频谱内选择半径 r_1 和 r_L 且满足 $r_1<r_L$，使 r_1 和 r_L 之间的环形区域覆盖中频带（为了取得鲁棒性和不可感知性的良好平衡，本文将数字水印嵌入到 DFT 域中频区）。设 $\{C(r_i),i=1,\cdots,L\}$ 是中频带内半径由小到大的同心圆族，且满足 $r_1\leqslant r_i\leqslant r_L$。

最后在 DFT 中频区域坐标轴的第一象限 E 内通过密钥 Key 产生一个伪随机点集 $\{(u_i,v_i),i=0,1,\cdots,P\times Q-1\}$，这些点是用来嵌入水印的位置。

（3）数字水印的嵌入

最后，对于点集中任意一点 $\{(u_i,v_i),i=0,1,\cdots,P\times Q-1\}$，选择点 $(-v_i,u_i)$ 与其配对，它们与中心点的连线成 90 度角。每一个 $C(r_i)$ 上都有这样一组点对 $((u_i,v_i),(-v_i,u_i))$，其中 $u_i^2+v_i^2=r_i^2$。对这些点对的幅值 $(M(u_i,v_i),M(-v_i,u_i))$ 进行量化调制，从而将 1bit 水印信息 $vm(i)$ 嵌入到 $C(r_i)$ 中。具体量化规则如下：

如果 $vm(i)=1$，则

如果 Δ 小于 α，则

$$M'(u_i,v_i)=M(u_i,v_i)+\left(\frac{\alpha}{2}-\frac{\Delta}{2}\right)$$

$$M'(-v_i,u_i)=M(-v_i,u_i)-\left(\frac{\alpha}{2}-\frac{\Delta}{2}\right) \tag{3}$$

否则

不做任何修改

如果 $vm(i)=0$,则

如果 Δ 大于 $-\alpha$,则

$$M'(u_i,v_i)=M(u_i,v_i)-\left(\frac{\alpha}{2}+\frac{\Delta}{2}\right)$$

$$M'(-v_i,u_i)=M(-v_i,u_i)+\left(\frac{\alpha}{2}+\frac{\Delta}{2}\right) \quad (4)$$

否则

不做任何修改

式中,$(M'(u_i,v_i),M'(-v_i,u_i))$ 为 $(M(u_i,v_i),M(-v_i,u_i))$ 的量化修改值;α 为阈值门限;

$$\Delta=M(u_i,v_i)-M(-v_i,u_i) \quad (5)$$

此外,为了保证修改结果的 IDFT 变换为实数,被修改点的中心对称点处的幅值也需进行相应修改。

(4)含水印图像获得。用修改后 DFT 系数代替原 DFT 系数并进行逆 DFT 变换,即可得到含水印图像 $\boldsymbol{F}'$。

3　数字水印的检测

本文结合 SVR 理论,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组合矩作为特征向量,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。

3.1　特征向量的构造

一般说来,几何攻击包括旋转、缩放、平移、剪切等多种形式。本文将重点讨论常见的旋转、缩放、平移等变换形式,显然该类变换在图像缩放纵横比为 1 时可用 4 个参数进行描述

$$\begin{bmatrix}x_n\\y_n\end{bmatrix}=\begin{bmatrix}s&0\\0&s\end{bmatrix}\cdot\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}\cdot\begin{bmatrix}x_0\\y_0\end{bmatrix}+\begin{bmatrix}t_x\\t_y\end{bmatrix} \quad (6)$$

式中,(x_0,y_0) 表示未受攻击的图像像素坐标;(x_n,y_n) 表示攻击后的图像像素坐标;(t_x,t_y) 表示平移距离;$\begin{bmatrix}s&0\\0&s\end{bmatrix}$ 为缩放矩阵;$\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}$ 为旋转矩阵。

假设图像是定义在整数值笛卡儿坐标网格上的实值函数 $f(x,y)$,其中 $0\leqslant x<M,0\leqslant y<N$,则分别把图像的几何矩 $m_{p,q}$ 和中心矩 $\mu_{p,q}$ 定义如下:

$$m_{p,q}=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}x^p y^q f(x,y)$$

$$\mu_{p,q}=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(x-\bar{x})^p(y-\bar{y})^q f(x,y)$$

这里,$\bar{x}=\frac{m_{1,0}}{m_{0,0}},\bar{y}=\frac{m_{0,1}}{m_{0,0}}$。

由数字图像相关理论知,低阶图像矩可很好地反映整幅图像模式,而高阶矩主要反映图像细节信息。考虑到本文重点讨论旋转、缩放、平移等整体变换形式,故以下采用几何矩和中心矩复合而成的 6 个特征来反映图像全局信息,并进一步将其作为 SVR 训练特征向量。

(1)图像旋转特征

$$f_1=\frac{2\theta}{\pi};\ f_2=\mu_{1,2}+\mu_{3,0};\ f_3=\mu_{0,3}+\mu_{2,1}$$

式中,

$$\theta=\arctan\left(\mu_{0,2}-\mu_{2,0}+\sqrt{(\mu_{0,2}-\mu_{2,0})^2+4\mu_{1,1}^2}\right)/2\mu_{1,1}$$

(2)图像平移特征

$$f_4=\frac{m_{1,0}}{m_{0,0}},\ f_5=\frac{m_{0,1}}{m_{0,0}} \quad (7)$$

(3)图像缩放特征

$$f_6=\left[(\mu_{2,0}+\mu_{0,2}+\sqrt{(\mu_{2,0}-\mu_{0,2})^2+4\mu_{1,1}^2})/2\times\left(\mu_{2,0}+\mu_{0,2}-\sqrt{(\mu_{2,0}-\mu_{0,2})^2+4\mu_{1,1}}\right)/2\right]^{\frac{1}{4}}$$

3.2　SVR 训练模型的获得

相对于神经网络而言,回归型支持向量机具有更强的泛化能力和学习能力。为了获得 SVR 训练模型,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 $\boldsymbol{F}$ 以产生 K 个训练样本图像 $\boldsymbol{F}^k(k=1,2,\cdots,K)$,然后计算出每个训练样本图像 $\boldsymbol{F}^k$ 的 6 个组合矩

$$(f_1^k,f_2^k,f_3^k,f_4^k,f_5^k,f_6^k)(k=1,2,\cdots,K)$$

并将其作为训练特征向量。同时,将相应的变换(平移、旋转和缩放)参数$(k=1,2,\cdots,K)$作为训练目标值,于是可得到训练样本

$$\boldsymbol{\Omega}_k=(f_1^k,f_2^k,f_3^k,f_4^k,f_5^k,f_6^k,t_x^k,t_y^k,s^k,\theta^k)\quad(k=1,2,\cdots,K)$$

由于训练样本的选取范围和数量直接影响到参数预测的精确度,所以在[-45,45]范围内随机选取了 50 个旋转参数训练样本,在[0.5,2]范围内随机选取了 50 个缩放参数训练样本,X 轴方向在[5,30]范围内随机选取 50 个平移参数训练样本, Y 轴方

向在[5,30]范围内随机选取 50 个平移参数训练样本。在一定范围内总共选取了 200 个参数作为训练样本。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,采用 4 个 SVR 并行结构构成 MIMO 系统,SVR 的结构为 6 个输入,核函数采用 RBF 径向基函数。进行训练学习,即可获得 SVR 训练模型。

3.3 待检测图像的 SVR 几何校正

为了有效提取数字水印,首先利用 SVR 训练模型对待检测图像进行数据预测,然后结合预测输出结果对待检测图像进行几何校正,以抵抗各种几何攻击。基于 SVR 的待检测图像校正过程为

(1)计算出待检测图像 $\tilde{\boldsymbol{F}}$ 的 6 个组合矩($\tilde{f}_1$,$\tilde{f}_2$,$\tilde{f}_3$,$\tilde{f}_4$,$\tilde{f}_5$,$\tilde{f}_6$),并将其作为训练特征向量。

(2)以 6 个组合矩特征($\tilde{f}_1$,$\tilde{f}_2$,$\tilde{f}_3$,$\tilde{f}_4$,$\tilde{f}_5$,$\tilde{f}_6$)作为输入向量,利用已经获得的 SVR 训练模型对输入向量进行数据预测,从而得到相应的输出向量值(即几何变换参数 $\tilde{t}_x$,$\tilde{t}_y$,$\tilde{s}$,$\tilde{\theta}$)。

(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 $\tilde{\boldsymbol{F}}$ 进行几何校正(即对待检测图像 $\tilde{\boldsymbol{F}}$ 实施几何变换的逆变换,如旋转角度逆变换、平移参数逆变换等),从而得到待检测图像 $\tilde{\boldsymbol{F}}$ 的校正结果 $\boldsymbol{F}^*$。

3.4 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 $\boldsymbol{F}^*$,则数字水印检测过程如下:

(1)对校正后待检测图像 $\boldsymbol{F}^*$ 实施 2 维 DFT 变换(变换原点为图像中心),得到中心化频谱,即

$$F^*(u,v)=DFT(\boldsymbol{F}^*)$$

(2)根据密钥 Key,在坐标轴第一象限的环形区域$[r_1,r_L]$内选择 $P\times Q$ 个 DFT 系数 $F^*(u_i,v_i)$($i=0,1,2,\cdots,P\times Q-1$),其相应幅值谱为 $M^*(u_i,v_i)$,则可按照如下规则提取水印信息

$$vm^*(i)=\begin{cases}1 & M^*(u_i,v_i)-M^*(-v_i,u_i)\geqslant 0\\ 0 & M^*(u_i,v_i)-M^*(-v_i,u_i)<0\end{cases} \tag{8}$$

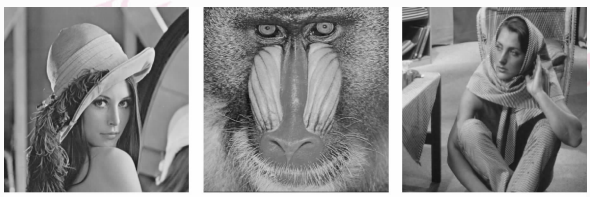
(3)对 $\boldsymbol{VM}^*$ 进行解密和升维处理(按照水印嵌入过程的相反操作),便可得到二值水印图像

$$\boldsymbol{W}^*=\{w^*(i,j),0\leqslant i<P,0\leqslant j<Q\}.$$

4 仿真实验结果

为了验证本文数字图像水印算法的有效性,以下分别给出了透明性测试、抗攻击能力(鲁棒性)测试的实验结果。实验中,所选用的原始载体分别为 512×512×8 bits 标准灰度图像 Lena、Mandrill 和 Barbara,数字水印采用了 32×32 的二值图像。内外半径分别为 $r_1=26$ 和 $r_L=86$,阈值门限为 $\alpha=1.2\times10^4$。训练样本数目为 $K=50$。SVR 训练时选用了 RBF 核函数。

图 1 为 Lena、Mandrill 和 Barbara 的含水印图像(利用本文算法)。



(a) Lena (PSNR=46.9 dB) (b) Mandrill (PSNR=44.0 dB) (c) Barbara (PSNR=46.3 dB)

图 1 数字水印的嵌入效果(本文算法)

Fig. 1 The watermarked image using our method

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)。

5 结 论

抗几何攻击的高度鲁棒数字图像水印算法研究是一项富有挑战性的工作。本文提出了一种可有效抵抗几何攻击的图像水印检测新算法,该算法首先选取图像的组合矩作为特征向量,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。仿真实验结果表明,本文算法对常规信号处理(滤波、叠加噪声、JPEG 压缩等)和几何攻击(旋转、缩放、平移、剪切等)均具有较好的鲁棒性。更为重要的是,本文所提出的 SVR 几何校正方法可与任

表 1　数字水印对常规信号处理的抵抗能力(失真率 BER)

Tab. 1　The watermark detection results for common signal processing

		未攻击	随机噪声	高斯噪声	椒盐噪声	均值滤波	高斯低通滤波	锐化
Lena	BER	0	0	0.04	0	0	0	0
	PSNR	46.90	33.39	25.21	30.51	31.81	39.90	22.71
Mandrill	BER	0	0	0.07	0	0	0	0.027
	PSNR	44.01	33.21	25.15	30.68	23.22	31.46	15.01
Barbara	BER	0	0	0.05	0	0	0	0
	PSNR	46.26	33.36	25.24	30.65	25.14	33.41	17.55

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)

Tab. 2　The watermark detection results for various attacks

		旋转 5°	旋转 45°	扩大 2 倍	缩小 0.5 倍	X 轴平移 5	X 平移 30	Y 轴平移 5
Lena	BER	0	0	0	0	0	0	0
	PSNR	12.09	8.63	12.66	12.69	18.15	12.11	19.73
Mandrill	BER	0	0	0	0.01	0	0	0
	PSNR	12.44	9.06	12.24	12.34	15.94	12.63	16.01
Barbara	BER	0	0	0	0.01	0	0	0
	PSNR	11.38	8.019	10.04	10.11	16.87	11.69	17.98

何其他水印嵌入方法结合,通用性较强。

此外,本文提出的 SVR 图像水印检测方案还具有计算简单、容易实现等特点,这大大增强了其用于数字图像作品版权保护的实用性,具有一定的应用价值。

参考文献(References)

1　Barni M, Cox I J, Kalker T. Digital watermarking [A]. In: Proceedings of the 4th International Workshop on Digital Watermarking 2005[C], Siena, Italy, 2005: 311-315.

2　Cox I J, Miller M L, Bloom J. Digital Watermarking [M]. San Francisco: Morgan Kaufmann Publishers, 2002.

3　Licks V, Jordan R. Geometric attacks on image watermarking system [J]. IEEE Multimedia, 2005, **1**(3): 68-78.

4　Ping D, Jovan G B, Nikolas P G. Digital watermarking robust to geometric distortions [J]. IEEE Transactions on Image Processing, 2005, **14**(12): 2140-2150.

5　Xin Y, Liao S, Pawlak M. A multibit geometrically robust image watermark based on zernike moments [A]. In: Proceedings of the 17th International Conference on Pattern Recognition [C], Cambridge, UK, 2004: 861-864.

6　Simitopoulos D, Koutsonanos D E. Robust image watermarking based on generalized radon transformations [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, **13**(8): 732-745.

7　Kim H S, Lee H K. Invariant image watermarking using zernike moments[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, **13**(8): 766-775

8　Pereira S, Pun T. Robust template matching for affine resistant image watermark [J]. IEEE Transactions on Image Processing, 2000, **9**(6): 1123-1129.

9　Kang Xiangui, Huang Jiwu, Shi Yun Q, *et al*. A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, **13**(8): 776-786

10　Tang C W, Hang H M. A feature-based robust digital image watermarking scheme [J]. IEEE Transactions on Signal Processing, 2003, **51**(4): 950-958

11　Jin S S, Chang D Y. Image watermarking based on invariant regions of scale-space representation [J]. IEEE Transactions on Signal Processing, 2006, **54**(4): 1537-1549.