文章编号:1671-6833(2009)02-0075-05

抗几何攻击的图像水印算法

李传目,宋海明,洪联系,万 春

(集美大学 计算机工程学院,福建 厦门 361021)

摘 要:提出了一种基于尺度不变特征变换(SIFT)和小波变换的抗几何攻击的自适应鲁棒水印算法.首先利用SIFT算法从载体图像中提取稳定的特征点;然后根据特征尺度和方向自适应来确定每个局部特征区域大小和方向;最后从中选择具有较大特征尺度互不重叠的特征区域,并利用量化小波系数的方法将水印嵌入到每个局部特征区域内.仿真实验结果表明,该算法不仅具有良好的透明性,而且具有较强的抵抗常规信号处理和几何攻击的能力.

关键词:数字水印;尺度不变特征变换;几何攻击;特征区域

中图分类号:TP 391 **文献标识码**:A

0 引言

目前大部分的数字水印方案能够抵抗常规信号处理,如滤波、增强、数据压缩等攻击,但对图像做一点微弱的几何变换(如旋转、缩放、平移等),就能够破坏水印检测的同步性,使得水印检测失败.近年来,研究者已经提出了许多抵抗几何攻击的水印方案[1-8],主要有以下几类[9]:①几何不变域方法.在对几何攻击具有不变性的变换域嵌入水印[1];②模板方法.在水印检测时,先利用模板估计几何变换的参数,然后进行相应的逆变换,再提取水印[2];③矩估计方法.通过几何矩或中心矩来估计水印图像所经历的几何变换参数,并利用此参数对图像进行校正,从而使水印检测过程和嵌入过程达到同步,以便进行正确的水印检测[3].如文献[3]中给出了一种基于原始图像矩进行几何失真估计的DCT域抗几何攻击算法;④基于特征点方法.在图像中检测稳定的特征点,图像在遭受各种几何攻击时,特征点基本不发生改变,从而确定水印的嵌入位置,实现水印的同步,如文献[4-6].这些方法能够抵抗较为一般的几何攻击,但有如下不足:鲁棒性较差;计算量较大;图像质量下降;稳定性差且分布不均匀等.

笔者以尺度不变特征变换(SIFT)为基础,结合小波变换相关知识,提出了一种可有效抵抗几何攻击的强鲁棒数字图像水印算法.

1 SIFT特征点的提取

SIFT算法是一种提取局部特征的算法,它在尺度空间寻找极值点,提取位置、尺度和旋转不变量.SIFT算法的主要步骤为:① 检测尺度空间极值点;② 精确定位极值点;③为每个特征点指定方向参数.尺度空间模拟图像数据的多尺度特征,一幅二维图像,在不同尺度下的尺度空间表示可由图像与高斯核卷积得到:

$$L(x,y,\sigma) = G(x,y,\sigma) \cdot I(x,y) \quad (1)$$

式中:(x,y)代表图像的像素位置;$G(x,y,\sigma)$是尺度可变高斯函数,其定义如下:

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

式中:σ称为尺度空间因子,其值越小则表征该图像被平滑的越少,相应的尺度也就越小;大尺度对应于图像的概貌特征,小尺度对应于图像的细节特征.

1.1 极值点坐标检测

满足在图像二维平面空间和DoG(Difference of Gaussian)尺度空间中同时具有局部极值的点作为SIFT特征点.DoG算子定义为两个不同尺度的高斯核的差分:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) \cdot I(x,y)$$

收稿日期:2008-11-09;修订日期:2008-12-21

基金项目:福建省自然科学基金资助项目(2008J0197,2006J0408);福建省教育厅科技项目(JA08139)

作者简介:李传目(1966-),男,河南范县人,集美大学副教授,主要研究方向为图像加密、数字水印等多媒体信息安全,E-mail:cm@jmu.edu.cn

$$= L(x,y,k\sigma) - L(x,y,\sigma) \quad (3)$$

通过计算某采样点在每一尺度下 DoG 算子的值,可以得到特征尺度轨迹曲线.特征尺度曲线的局部极值点即为该采样点的尺度.

1.2 极值点精确定位

通过拟和三维二次函数确定了特征点的位置和尺度,然而因为 DoG 算子会产生较强的边缘响应,所以 SIFT 算法需要舍弃低对比度的特征点和不稳定的边缘响应点以增强稳定性和提高抗噪声能力.舍弃特征点的依据是:一个定义不好的 DoG 的极值在横跨边缘的地方有较大的主曲率,而在垂直边缘的方向有较小的主曲率.而 DoG 的主曲率通过一个 2×2 的 Hessian 矩阵 $\boldsymbol{H}$ 求出:

$$\boldsymbol{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (4)$$

这里 $D_{xx}(D_{xy},D_{yy})$ 为 $x(x$ 和 $y,y)$ 方向的二阶导数. DoG 的主曲率和 $\boldsymbol{H}$ 的特征值成正比,令 α 为矩阵 $\boldsymbol{H}$ 的最大特征值,β 为矩阵 $\boldsymbol{H}$ 的最小特征值,则矩阵 $\boldsymbol{H}$ 的迹和行列式分别为:

$\mathrm{Tr}(\boldsymbol{H}) = D_{xx} + D_{yy} = \alpha + \beta$ 和 $\mathrm{Det}(\boldsymbol{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta$,令 $\alpha = \gamma\beta$,则有

$$\frac{\mathrm{Tr}(\boldsymbol{H})^2}{\mathrm{Det}(\boldsymbol{H})} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(\gamma\beta+\beta)^2}{\gamma\beta^2} = \frac{(\gamma+1)^2}{\gamma} \quad (5)$$

因为 $\frac{(\gamma+1)^2}{\gamma}$ 的值在两个特征值相等的时候最小,随着 γ 的增大而增大,所以检测主曲率是否在某域值 γ 下,只需检测 $\frac{\mathrm{Tr}(\boldsymbol{H})^2}{\mathrm{Det}(\boldsymbol{H})} < \frac{(\gamma+1)^2}{\gamma}$.一般 $\gamma = 10$.

1.3 特征点方向和尺度的性质

SIFT 算法利用特征点邻域像素的梯度方向分布特性为每个特征点指定方向参数,使算子具备旋转不变性.

$$m(x,y) = \sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2} \quad (6)$$

$$\theta(x,y) = \mathrm{atan}(L(x,y+1) - L(x,y-1)/L(x+1,y) - L(x-1,y) \quad (7)$$

式中:L 的尺度为每个特征点各自所在的尺度,则 $m(x,y)$ 和 $\theta(x,y)$ 为 (x,y) 处梯度的模值和方向.

对图像进行的缩放直接改变了图像的尺度,而 SIFT 特征点的获取是在图像的不同尺度下分别进行的,因而特征点的尺度特征与图像的缩放具有比例关系,假定缩放前特征点 F 的尺度为 σ_1,缩放后与之匹配的特征点尺度为 σ_2,且水印图像的缩放比例为 s,则由尺度空间的性质有:$\sigma_2 = s\sigma_1$[9].

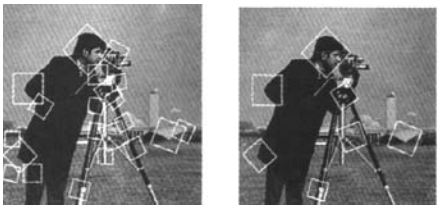
2 特征区域的选取

以图像特征点为标识,从载体图像中选出的一部分子图像,作为局部特征区域,用来嵌入数字水印区域,由于特征尺度仅取决于图像局部特性,其大小随图像局部特性改变而改变,且大尺度对应于图像的概貌特征,小尺度对应于图像的细节特征,这样特征尺度较大的特征点在经过常规图像处理后具有较好的稳定性,因此可以以其为中心选择局部特征区域.首先将坐标轴旋转为特征点的方向,以确保旋转不变性;然后以特征点为中心,以其所对应的特征尺度倍数作为边长,从载体图像中自适应地选出正方形的局部特征区域,而其边长大小定义为:

$$R = k \cdot \mathrm{round}(\sigma) \quad (8)$$

式中:R 表示局部特征区域边长;σ 表示当前图像特征点的特征尺度;round(·)为四舍五入取整函数;k 为常正整数,用于调节 R 的大小,即局部特征区域大小.

特征区域可能存在嵌套或重叠的情况.为了保证特征区域没有嵌套或重叠,首先设定一个特征尺度阈值 $\hat{T}$,仅保留特征尺度大于域值的稳定的特征区域,其他舍弃(如图 1(a)所示).其次对于嵌套或重叠的每一组特征区域,仅保留特征尺度最大(稳定性最好)的特征区域,并除去与其存在重叠的其它特征区域内的特征区域(如图 1(b)所示).



(a)特征区域 (b)不重叠特征区域

图 1 特征区域

Fig.1 feature regions

3 水印的嵌入

设原始图像为 I,数字水印的嵌入过程关键步骤如下:

(1)水印产生.由密钥 Key_1 产生一个伪随机序列 $W = \{w_i, i = 1,\cdots,L\}$ 作为数字水印,其中,L

为水印长度，$w_i \in \{0,1\}$.

(2)利用高斯滤波器对图像 I 进行平滑处理，以消除噪声干扰；

(3)局部特征区域. 利用 SIFT 算法，按照第 3 节所述方法，在图像 I 中选择一组特征尺度较大的稳定的局部特征点及相应互不重合的正方形特征区域 $O=\{o_k,k=1,\cdots,m\}$；

(4)小波变换. 对每个方形局部特征区域实施 n 级小波变换，并选取低频子带 LL_n 作为数字水印嵌入区(每个低频子带都嵌入相同的水印)；

(5)选择小波系数. 利用密钥 Key_2 在低频子带 LL_n 内随机选取不重合的 L 个位置作为水印信号的嵌入位置，记作 $C=\{(x_k,x_k),k=1,\cdots,L\}$.

(6)量化嵌入. 采用下列方法进行量化嵌入水印：

$$Q(LL_n(x_k,x_k))=\mathrm{floor}(\mathrm{round}(LL_n(x_k,x_k)/\Delta(x_k,x_k))/2)\times 2\,\Delta(x_k,x_k) \tag{9}$$

$$\tilde{LL}_n(x_k,x_k)=Q(LL_n(x_k,x_k))+w_k\cdot\Delta(x_k,x_k) \tag{10}$$

式中：$LL_n(x_k,x_k)$ 和 $\tilde{LL}_n(x_k,x_k)$ 分别为修改前后的小波系数；$Q(\cdot)$ 为量化器；floor $(\cdot)$ 为地板函数；$\Delta(x_k,x_k)$ 为量化步长，因为量化步长与水印嵌入强度密切相关，取值越大，数字水印鲁棒性能越好，但同时图像的视觉效果较差. 由于图像的最终接收者为人，因此量化步长的选取应充分考虑图像自身特点和人眼视觉特性，量化步长取值为：

$$\Delta(x_k,y_k)=\alpha\times 2^n\times \ln\frac{\left|LH_n(x_k,y_k)\right|+\left|HL_n(x_k,y_k)\right|+\left|HH_n(x_k,y_k)\right|}{2} \tag{11}$$

式中：α 为拉伸因子.

(7)逆小波变换. 用含水印信息的小波系数 $\tilde{LL}_n(x_k,x_k)$ 代替 $LL_n(x_k,x_k)$，并结合未修改的其他小波系数进行 n 级逆小波变换，便可得到含水印图像块，用其替换相应的局部特征区域. 对每个局部特征区域都做上述处理，最后即得到含水印图像 $\tilde{I}$.

4　水印的提取

水印的提取为嵌入的逆过程，首先按照嵌入过程中的方法选择若干个互不重叠的稳定的局部特征区域，且只要从两个局部特征区域能够检测到水印，便可认为数字水印存在于待检测图像中. 设待提取水印的图像为 $\tilde{I}$，整个数字水印提取过程的主要步骤如下：

(1) 对待检测图像 $\tilde{I}$ 使用相同密钥 Key_2，进行同第 3 节嵌入过程中的步骤(2)～(5)，可以在低频子带 LL_n 内选取不重合的 L 个位置用来提取水印，记作 $C=\{(x_k,x_k),k=1,\cdots,L\}$.

(2) 同样，用第 3 节嵌入过程中步骤(6)中的量化方法及量化步长 $\Delta(x_k,x_k)$，按下式提取水印：

$$\hat{w}_k=\mathrm{mod}(\mathrm{round}(LL_n(x_k,x_k)/\Delta(x_k,x_k)),2) \tag{12}$$

式中：mod $(\cdot,2)$ 为模 2 取余操作.

这样，就可以提取水印 $\hat{W}=\{\hat{w}_i,i=1,\cdots,L\}$.

(3) 在数字水印检测过程中，经常会发生虚警错误(false, alarm, error). 为了降低虚警错误的概率(即虚警率)，需要对所提取出的水印与原始数字水印进行比较，当匹配的比特数大于某设定阈值时，才能认为待检测图像中存在数字水印.

为此，计算提取水印 $\hat{W}$ 和原始水印 W 的归一化汉明相似度(NHS)：

$$\mathrm{NHS}=1-\mathrm{HD}(\hat{W},W)/L \tag{13}$$

其中，$\mathrm{HD}(\hat{W},W)$ 表示水印系列 $\hat{W}$ 和 W 之间的汉明距离.

将水印序列的每一位看作独立变量，对长度为 L 比特的水印，提取的水印中有 k 比特与原始水印正确匹配的概率为：

$$P_k=\binom{L}{k}p^k(1-p)^{L-k} \tag{14}$$

式中：P_k 表示提取水印的每一位与原始水印的对应位正确匹配的概率，由于水印位为均匀分布的二值水印，所以 $p=0.5$.

于是，每一个特征区域中提取水印的虚警率与设定的阈值 T 之间的关系为：

$$P_{\mathrm{local}}=\sum_{k=T}^{L}0.5^L\frac{L!}{k!(L-k)!} \tag{15}$$

式中：k 表示提取水印与原始水印正确匹配的位数.

在本文中，若能从两个特征区域中成功提取水印，则表明水印存在. 因此，从图像中提取水印的虚警率为：

$$P_{\mathrm{global}}=\sum_{i=2}^{m}(P_{\mathrm{local}})^i(1-P_{\mathrm{local}})^{m-i}\binom{m}{i} \tag{16}$$

式中：m 表示特征区域的个数. 当水印长度为 64 比特(本文中水印长度)，水印检测阈值 T 为 50(本文中的阈值)，对应的 NHS 为 0.781，虚警率为 10^{-4}.

5 仿真实验

为了验证本文数字图像水印算法的高效性,分别进行了不可见性测试、常规信号处理及抗几何攻击测试. 在实验中,如图2所示,所选用的原始载体分别为512×512×8b标准测试图像'Cameraman','Baboon','Peppers'. 因为水印长度太短,不足以说明图像中水印的存在;水印长度太长,将会增加计算量,并且降低图像质量,所以水印长度选择64位(检测阈值$T=50$).

(a) Cameraman (b) Baboon (c) Peppers

图2 原始图像及特征区域

Fig. 2 Original Images and feature regions

5.1 不可见性

按本文方法嵌入水印,嵌入水印后的图像如图3所示,表1给出了嵌入水印后的图像不可见性测试结果. 根据实验结果可以看出,含水印图像与原始载体间的峰值信噪比远大于40 dB,从人的视觉方面,根本看不出它们之间的差别,说明嵌入水印后,图像的质量没有明显下降.

(a) Cameraman (b) Baboon (c) Peppers

图3 含水印图像

Fig. 3 Watermarked Images

表1 含水印图像与原始载体间的PSNR

Tab. 1 PSNR Between Watermarked Image dB

图像	Cameraman	Baboon	Peppers
PSNR	48.76	47.80	48.54

5.2 常规图像处理

为了检测本文算法的鲁棒性能,分别对含水印图像进行了一系列常规图像处理仿真实验,包括锐化、加噪、JPEG压缩及其混合处理,然后再提取嵌入的水印,表2给出了本文算法实验结果及与其他方法的结果对比,表2中分母表示从受攻击图像中检测出的特征区域数目,分子表示能成功提取水印的特征区域数目,×表示没有给出实验结果.

表2 常规信号处理后正确提取水印的特征区域数

Tab. 2 Number of feature regions extracting correct watermarking under common signal processing

Attack	Cameraman	Baboon		Peppers	
	本文算法	本文算法	文献[6]算法	本文算法	文献[6]算法
Sharpening(3×3)	6/7	6/11	4/11	7/9	4/4
Salt & Pepper noise(0.02)	6/7	5/11	×	6/9	×
Median filter (3×3)	5/7	8/11	×	7/9	×
Gaussian filter (3×3)	4/7	10/11	8/11	5/9	1/4
JPEG_30	3/7	6/11	4/11	4/9	0/4
JPEG_90	6/7	11/11	×	6/9	×
JPEG_90 + Sharpening(3×3)	5/7	6/11	2/11	8/9	4/4
JPEG_90 + Gaussian filter(3×3)	4/7	10/11	8/11	7/9	2/4

5.3 几何攻击

为了检测本文算法的抗几何攻击能力,分别对含水印图像进行了一系列几何攻击仿真实验,包括删除行列、剪切、加噪、shearing、缩放、旋转等攻击,表3给出了本文算法实验结果及与其他方法的结果对比,表中分母表示从受攻击图像中检测出的特征区域数目,分子表示能成功提取水印的特征区域数目,×表示没有给出实验结果.

从实验结果中可以看出,本文算法对于常规图象处理及旋转、缩放等几何攻击及其组合攻击具有很好的鲁棒性. 对实验中测试的各种几何攻击,算法均能够成功地提取水印,所有情况下成功提取水印的特征区域数目都大于2个,正确提取水印的比率高于其它同类算法.

6 结论

笔者以SIFT算法为基础,提出了一种可有效抵抗几何攻击的强鲁棒数字图像水印算法,具有以下主要特点:①所提取的图像特征点不仅稳定性好,而且分布均匀. 由于选择具有较大尺度特征

的特征点,而大尺度对应于图像的概貌,所以提高　了整个水印系统对常规信号处理、随机剪切等攻击的抵抗能力;②能够结合图像内容自适应确定局部特征区域大小,由特征尺度的性质知道,尺度特征与图像的缩放具有比例关系,所以能有效抵抗缩放等仿射变换攻击;③由 SIFT 算法的特点知道,算法简单、容易实现,图像旋转前后确定的特征区域不变,所以能有效抵抗旋转攻击;且抽取水印时无需原始载体,改善了整个系统的抗局部攻击能力.

表 3　遭受几何攻击后正确提取水印的特征区域数

Tab. 3　Number of feature regions extracting correct watermarking under geometric attacks

Attack	Cameraman	Baboon		Peppers	
	本文算法	本文算法	文献[6]算法	本文算法	文献[6]算法
17_row_5_col_removed	6/7	7/11	5/11	6/9	1/4
scale_0.90	6/7	8/11	×	6/9	×
scale_1.10	6/7	8/11	×	6/9	×
rotation_30.00	6/7	9/11	×	7/9	×
Cropping_10 + rotation_5	5/7	7/11	0/11	5/9	0/4
JPEG70 + 17_row_5_col_removed	5/7	10/11	3/11	8/9	1/4
JPEG70 + rotation_30.00	5/7	7/11	×	6/9	×
JPEG70 + cropping_10	6/7	8/11	×	7/9	×
JPEG70 + scale_0.90	5/7	6/11	×	5/9	×
JPEG70 + scale_1.10	5/7	6/11	×	5/9	×

参考文献:

[1] ZHENG D, ZHAO J Y. RST invariant digital image watermarking: Importance of phase information [C]//IEEE Canadian Conference on Electrical and Computer Engineering. Montreal, Canada: IEEE Press, 2003: 785 - 788.

[2] QI X J, QI J. Improved affine resistant watermarking by using robust templates [C]//IEEE International Conference on Acoustics, Speech, and Signal Processing. Montreal, Canada: IEEE Press, 2004: 405 - 408.

[3] 张　力,肖薇薇,张基宏. 基于原始图像矩的仿射不变性水印算法[J]. 深圳大学学报,2003,20(2):16 - 21.

[4] TANG C W, HANG H M. A feature - based robust digital image watermarking scheme [J]. IEEE Trans on Signal Processing ,2003, 51(4): 950 - 958.

[5] LOWE D G. Distinctive image features from scale - invariant keypoints[J]. Interna tional Journal of Computer Vision, 2004, 60(2): 91 - 110.

[6] 徐振启,卢　洵,罗少鹏. 基于时空二维混沌的自适应彩色图像水印算法[J]. 郑州大学学报:工学版, 2007, 28(2):84 - 87.

[7] 王向阳,侯丽敏,杨红颖. 基于图像特征点与伪 Zernike 矩的鲁棒水印算法研究[J]. 计算机研究与发展, 2008, 45(5):772 - 778.

[8] 李振宏,吴慧中. 基于 SIFT 变换的水印图像几何失真校正算法[J]. 计算机工程与设计, 2008, 29(12):3215 - 3217.

An Image Watermarking Algorithm Robust to Geometric Attacks

LI Chuan - mu, SONG Hai - ming, HONG Lian - xi, WAN Chun

(School of Computer Engineering, Jimei University, Xiamen 361021, China)

Abstract: A geometrically robust image - adaptive watermarking scheme based on scale invariant feature transform (SIFT) and DWT is proposed. Firstly, the steady feature points are extracted from the host image by using SIFT; then the size and orientation of each local feature region are determined adaptively according to its feature scale and orientation. Lastly, some unoverlapped local feature regions with large - scale are selected for embedding watermarking, and the same digital watermarking is embedded into each local feature region by quantizing its wavelet coefficients. Simulation results show that the proposed watermaking scheme is not only invisible, but also robust against common signal processing and geometric attacks.

Key words: digital watermarking; SIFT; geometric attacks; feature region